UNIVERSITY OF COPENHAGEN

This thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen.

From Participation Factors to Co-Calibration of Patient- and Wearable-Reported Outcomes in Behavioural, Health, and Quality of Life Studies

PhD thesis

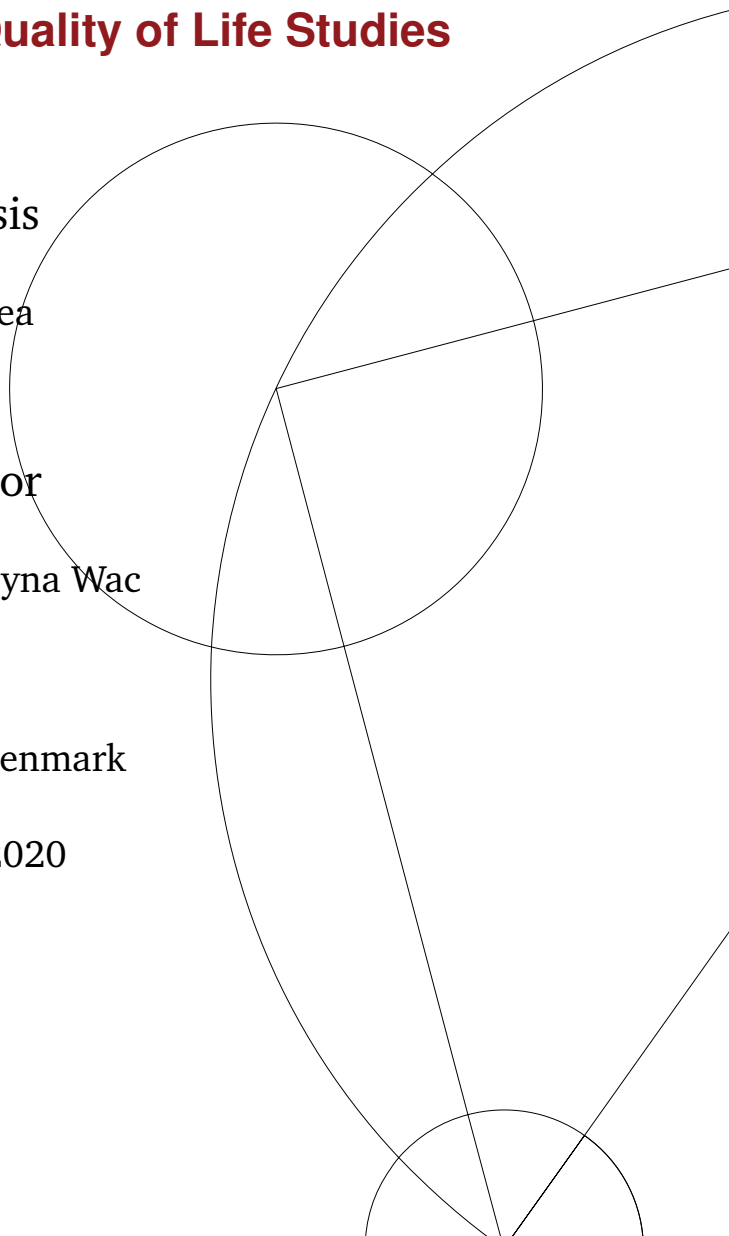
Vlad Manea

Supervisor

Professor Katarzyna Wac

Copenhagen, Denmark

December 2020



University of Copenhagen | Københavns Universitet

Faculty of Science
Department of Computer Science
Section of Human-Centered Computing
Quality of Life Technologies Lab

From Participation Factors to Co-Calibration of Patient- and Wearable-Reported Outcomes in Behavioural, Health, and Quality of Life Studies

Vlad Manea
PhD thesis

Supervisor Professor Katarzyna Wac
Université de Genève, Geneva, Switzerland
Københavns Universitet, Copenhagen, Denmark

Reviewer Professor Nancy E. Mayo
McGill University, Montreal, Canada

Reviewer Professor Bert Arnrich
Universität Potsdam, Potsdam, Germany

Reviewer Professor Thomas T. Hildebrandt
Københavns Universitet, Copenhagen, Denmark

This thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen.

From Participation Factors to Co-Calibration of Patient- and Wearable-Reported Outcomes in Behavioural, Health, and Quality of Life Studies

Vlad Manea, PhD thesis, Copenhagen, Denmark, December 2020

Supervisor, Professor Katarzyna Wac

University of Copenhagen | Københavns Universitet

Faculty of Science

Department of Computer Science

Section of Human-Centered Computing

Quality of Life Technologies Lab

Sigurdsgade 41, 2200, Copenhagen, Denmark

Abstract

Chronic diseases represent a significant share of the burden of disease globally. They are responsible for 86% of premature deaths in Europe. Unhealthy behaviours, such as physical inactivity, insufficient sleep, poor nutrition, and tobacco intake, explain up to 50% of chronic disease risk. However, the evidence is not precise enough to assess the risk for each disease. Human subject studies monitoring behaviours over long periods (longitudinally) during daily life (in situ) by leveraging unobtrusive (observational) technology can allow human behaviours to unfold. They can not only qualify, but also quantify the relationships between behaviours, health, and Quality of Life (QoL) outcomes from compliant participants.

This PhD thesis explores two research areas. In the first area, we research the motivation and facilitation of participation in human subject studies. We propose a presentational model using personalised stories to improve human studies' participation. We design two unifying frameworks for conducting a wide range of human subject studies (mQoL mobile app, mQoL-Chat chatbot). They leverage two modules designed and developed by the author in mQoL-Lab, the lab platform of the Quality of Life Technologies lab.

In the second area, we research the relationships between behavioural, health, and QoL outcomes (co-calibration). We present the coQoL computational model for co-calibration. We demonstrate its feasibility in a study on N = 42 healthy older individuals (a population at risk, appropriate for disease prevention, and having benefitted from insufficient co-calibrations). They answered questionnaires on eight physical and psychological validated scales (physical activity: IPAQ, social support: MSPSS, anxiety and depression: GADS, nutrition: PREDIMED and SelfMNA, memory: MFE, sleep: PSQI, and health-related QoL: EQ-5D-3L). They wore consumer wearables (Fitbit Charge 2) for up to two years. The wearables reported behavioural markers (physical activity, sleep, heart rate) in situ. We observed new relationships between these outcomes. We described the study's human factors and data quality.

The scientific contributions in both research areas can inform the design of future studies leveraging consumer technology that monitors behaviours longitudinally in situ to assess and improve health and QoL.

Résumé

Kroniske sygdomme udgør en betydelig del af sygdomsbyrden globalt. De er ansvar lige for 86% af de tidlige dødsfald i Europa. Usunde vaner, såsom fysisk inaktivitet, utilstrækkelig søvn, dårlig ernæring og tobaksindtagelse, er skyld i op til 50% af risikoen for kronisk sygdom. Men beviserne er ikke præcise nok til at vurdere risikoen for hver sygdom. Undersøgelser, der overvågner adfærd over længere perioder (langsgående) i dagligdagen (in situ) ved at udnytte diskret (observationel) teknologi, kan lade individets adfærd udfolde sig. De kan ikke kun kvalificere, men også kvantificere forholdet mellem adfærd, sundhed og livskvalitet (QoL) fra deltagere, der overholder reglerne.

Denne ph.d.-afhandling undersøger to forskningsområder. Pådet første område forsker vi i motivationen og faciliteringen af deltagelse i undersøgelser. Vi foreslår en præsentationsmodel, der bruger personliggjorte historier for at forbedre deltagelse i undersøgelser. Vi designer to forendende ramerværker for at gennemføre en lang række undersøgelser (mQoL-mobil app, mQoL-Chat chatbot). De udnytter moduler designet og udviklet af forfatteren i mQoL-Lab, labplatformen af Quality of Life Technologies Lab.

Pådet andet område forsker vi forholdet mellem adfærdsmæssige, sundhedsmæssige og QoL-resultater (co-kalibrering). Vi præsenterer coQoL beregningsmodel for co-kalibrering. Vi demonstrerer coQoLs gennemførlighed i en undersøgelse af N = 42 raske ældre individer (de er en risikogruppe, der passer til sygdomsforebyggelse, og som ikke har haft fordel af nok co-kalibrering). De svarede påspørgeskemaer, der havde otte fysiske og psykologiske valideringsskalaer (fysisk aktivitet: IPAQ, social støtte: MSPSS, angst og depression: GADS, ernæring: PREDIMED og SelfMNA, hukommelse: MFE, søvn: PSQI og sundhedsrelateret QoL: EQ-5D-3L). De havde brugt armbåndsure (Fitbit Charge 2) i op til to år. De armbåndsurene reporterede adfærdsmarkører (fysisk aktivitet, søvn, hjerterytme) in situ. Vi observerede nye forhold mellem disse resultater. Vi beskrev undersøgelsens menneskelige faktorer og datakvalitet.

De videnskabelige bidrag i begge forskningsområder kan bruges til at designe fremtidige undersøgelser, der udnytter forbrugerteknologi, som overvåger adfærd langsgående og in situ for at vurdere og forbedre sundheden og QoL.

Preface

This thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen. I have carried out the PhD research between January 2018 and December 2020 at the Quality of Life Technologies Lab (*QoL Lab*).

I thank my supervisor, professor Katarzyna Wac, for her support, endurance, and leadership. I am grateful to professors Tariq Andersen and Jonas Moll for their warm welcome at Vital Beats in Fall 2020. Furthermore, I thank professors Vero Estrada Galiñanes, Naja Holten Møller, and Nancy E. Mayo as well as colleagues Allan Berrocal, Elisabeth Brinth Refstrup, Alexandre De Masi, Semahat Ece Elbeyi, Jamshed Gill, Alba Kejser Perez, Kirke Kjellberg, Sofia Laghouila, Kedar Mate, Cecilie Rosentoft, Nadine Sandjo, Mads Schnoor Hansen, and Natalie Solomon for our collaborations during my PhD. I thank the participants in the CoME project for their contributions. I thank my friends Taregh Jasemian, Florin Pogocsan, and Adrian Popovici for their help with the thesis' introduction. I thank professors Nancy E. Mayo, Bert Arnrich, and Thomas T. Hildebrandtfor reviewing the thesis.

Acknowledgements are due to professors Lenuţa Alboaie, Stephen Alstrup, Sabin Buraga, Dan Cristea, Corina Forăscu, Cristian Gaţu, Florin Iacob, Adrian Iftene, Henri Luchian, Troels Mønsted, and Ferucio Laurenţiu Ţiplea for our collaborations, discussions, or help before my PhD. In particular, I thank professors Emanuela Cerchez and Marinel Şerban for their top-quality education, personal support, and impactful collaborations. I am also grateful to professors Mirela Cerchez, Rodica Cricler, Ioan Milică, Luiza Potoroacă, Mihaela Tudose, and Corina Stancu for their personal time investment that contributed to my early development. I thank professor Monika Marcu for her top-quality education in the earliest of times.

I could not have written this thesis without Filipa and her love. I also thank my family for their care over time. Also, I am fortunate to have very strong friendships over the years from Paul Diac, Codrin Dițu, Bogdan Florescu, Florin Pogocsan, Adrian Popovici, the public speaking and leadership community in Copenhagen, the med-tech innovators in Copenhagen, the algorithms and programming group in Iași, the algebraic foundations group in Iași, the trekking group in Iași, and more. :)

Vlad Manea
Copenhagen, Denmark | December 2020

Contents

1 Introduction

Chronic diseases represent a significant share of the burden of disease globally [1]. They are responsible for 86% of premature deaths in Europe [2]. Unhealthy behaviours, such as physical inactivity, insufficient sleep, poor nutrition, and tobacco intake, explain up to 50% of chronic disease risk [3]. For example, "*there is overwhelming evidence that proves the notion that reductions in daily physical activity are a primary cause of chronic diseases*" [4]. While the evidence is "*strong enough to cover all health outcomes*", it is "*currently insufficiently precise to warrant separate guidelines for each specific disease*" [5]. Regarding health, researchers assess the risk of cardiovascular disease, the most prevalent chronic disease and the primary cause of mortality worldwide [1], in cohorts followed over the years to tens of years (*longitudinal*) by using qualitative methods [6]. Studies monitoring behaviours longitudinally in the context of daily life (*in situ*) by leveraging unobtrusive (*observational*) technology can qualify and quantify the impact of behaviours on health and Quality of Life (*QoL*). We denote them as *human subject studies*.

Research Area 1: Motivation and Facilitation of Human Subject Study Participation

Numerous factors challenge the motivation to participate in human subject studies even before assessing behaviours. However, researchers focused more on participants' eligibility criteria than motivation to participate. They traditionally identified the motivation factors by surveying the individuals on their motivation to participate or reported as limitations *retrospectively*. Participant attrition occurred after only days to weeks, allowing only *momentary* assessments. Furthermore, the responses may have suffered from the inherent biases of self-reporting [7, 8]. Instead, participation can be motivated, assessed, and reported *prospectively* in situ and longitudinally by using mobile and wearable technologies, from the moment of enrollment, throughout the study, and up to the study's abandonment or completion (whichever occurs first).

The **motivation and facilitation of human subject study participation** is the first area explored in this thesis. We reviewed the factors, challenges, and opportunities

to participate in human subject studies and, in some instances, *longitudinal* studies or *health* studies. We produced a presentational model that uses *personalised stories* to motivate participation. We designed two unifying frameworks to facilitate participation: the *mQoL* mobile app design (2018) and the *mQoL-Chat* chatbot design (2019). Finally, we designed and developed modules in the Mobile Quality of Life platform (*mQoL-Lab*) used for human subject studies in the Quality of Life Technologies Lab (*QoL Lab*).

Research Area 2: Co-Calibration of Behavioural, Health, and Quality of Life Outcomes

Once individuals participate in research, a study can assess behaviours, health, and QoL, by using a combination of *reported outcomes*: patient-reported (PRO, [9]), performance-reported (PerfRO, [9]) technology-reported (TechRO, [9]), and in fewer instances other types.

PROs refer to questionnaires with validated scales that assess individual outcomes *momentarily* or for a given *recall* period (e.g., two weeks). PerfROs refer to physical or mental exercises/tests/protocols that assess momentary states and performance (maximum capacity) of the individual (e.g., the six-minute walk test [10]). PROs and TechROs are currently the scientific *gold standard* in assessing behaviours and health. However, PROs and PerfROs are *inconvenient* (implying participant deliberate effort) and *infrequent* (usually, coinciding with doctor visits). Furthermore, PROs are *recalled* (selected by participant memory), *socially conditioned* (participants may give socially acceptable answers to avoid judgement), *subjective* (perceived instead of actual), and ultimately *qualitative* [11].

Meanwhile, TechROs provided by *emerging* and increasingly accurate wearables are *frequent* (down to the millisecond), *consistent* (collected immediately and persisted for future use, yet subject to removal), *non-judgemental* (not prone to socially acceptable answers), *objective* (actual, yet subject to measurement accuracy), and *quantitative* [11]. We call such TechROs *digital biomarkers* [12] (or digital *behavioural markers* [13]). Digital biomarkers can estimate behaviours with enough accuracy to be leveraged in human subject studies, e.g., physical activity (e.g., energy expenditure, steps, distance, elevation), sleep (e.g., duration), body temperature, respiration rate, heart rate, perspiration (e.g., galvanic skin response).

In Europe, chronic diseases affect over 80% of adults over 65 and incur 70% of the increasing healthcare costs [14]. The importance of the long-term assessment

of behaviours to quantify health and risk of disease in the future is increasing as the world population is ageing [15], and age dramatically contributes to the risk of multiple diseases [1] in itself. Therefore, the healthy old (*seniors*) are a population both inherently at risk and appropriate for primary disease prevention. However, little research focused on the unobtrusive quantification of the relationships (*co-calibration*) between in situ longitudinal behaviours, health, and QoL.

Prior work in the co-calibration of PROs with TechRO focused on specific PROs suitable for the study aim; some PROs are disease-specific, which also relate to the study's user groups (e.g., students, adults with a given condition). As for the TechROs, we observed few research wearables (validated, expensive, and bulky lab-grade devices, used for a limited time — usually, under one month — due to the user burden and discomfort of wearing them), and several consumer wearables (e.g., Fitbit, Withings, Apple Watch, mostly worn as fitness bracelets).

Given state of the art on co-calibration, the second research area explored in this thesis is the **co-calibration of behavioural, health, and QoL outcomes** in healthy seniors using momentary PROs from questionnaires with validated scales, and longitudinal TechROs from consumer wearables. We produced the *coQoL* computational model for the co-calibration of PROs and TechROs. We applied coQoL in an observational, longitudinal, and in situ human subject study on seniors that collected empirical data on physical and psychological PROs, and digital biomarker TechROs. We first assessed the data quality. Then, we reported PRO-TechRO patterns of relationships by using coQoL.

We also included in the thesis four modules designed and developed during this PhD. Two modules (*questionnaire data collection* and *consumer wearable data collection*) are additions to the mQoL-Lab and serve as the tools providing the PROs and TechROs in the study applying coQoL.

Thesis Structure

The thesis is organised as follows. The first part is the introduction (Part 1). The second part presents research on *the motivation and facilitation of human subject study participation* (Chapter 2). The third part describes research on *the co-calibration of behavioural, health, and QoL outcomes* (Chapter 3). The fourth part summarises *the design and development of the mQoL-Lab platform* (Chapter 4). The fifth part concludes the thesis and provides areas of future work (Part 5).

List of Publications

Motivation and Facilitation of Human Subject Study Participation

1. Vlad Manea, Mads Schnoor Hansen, Semahat Ece Elbeyi, Katarzyna Wac. *Towards Personalizing Participation in Health Studies*. Workshop on Multimedia for Personal Health and Health Care (HealthMedia 2019), Conference on Multimedia (MM 2019). 8p. DOI: `https://doi.org/10.1145/3347444.3356241`.

2. Vlad Manea, Katarzyna Wac. *mQoL: Mobile Quality of Life lab: from Behavior Change to Quality of Life*. Workshop on Mobile Human Contributions (MHC 2018), Conference on Pervasive and Ubiquitous Computing (UbiComp 2018). 6p. DOI: `https://doi.org/10.1145/3267305.3267549`.

3. Vlad Manea, Allan Berrocal, Alexandre De Masi, Naja Holten Møller, Katarzyna Wac, Hannah Bayer, Sune Lehmann, Euan Ashley. *Call for Papers: LDC '19: Workshop on Longitudinal Data Collection in Human Subject Studies*. Call for Papers for the Workshop on Longitudinal Data Collection in Human Subject Studies (LDC 2019), Conference on Pervasive and Ubiquitous Computing (UbiComp 2019). 4p. DOI: `https://doi.org/10.1145/3341162.3347758`.

4. Allan Berrocal, Vlad Manea, Alexandre De Masi, Katarzyna Wac. *mQoL-Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices*. Conference on Mobile Systems and Pervasive Computing (MobiSPC 2020). 9p. DOI: `https://doi.org/10.1016/j.procs.2020.07.033`. **Nominated for the best project pitch award at the University Hospitals of Geneva Innovation Day 2020, Geneva, Switzerland**.

5. Vlad Manea, Vero Estrada-Galiñanes, Katarzyna Wac. *mQoL: Mobile Quality of Life lab*. Poster and demo at the Digital Health Conference (DH 2018). **Nominated for the Innovation Prize in the category of the best data-driven innovation, Lyon, France**.

Co-Calibration of Behavioural, Health, and Quality of Life Outcomes

6. Vlad Manea, Katarzyna Wac. *Co-calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method*. Journal of Personalized Medicine, 10(4), MDPI, 2020. Special Issue: *PROomics: Patient Reported Outcome (PRO) and Self-Tracking for Personalized Medicine*. Impact factor 4.433, rank 10/102 (Q1) in Health Care Sciences and Services. 41p. DOI: `https://doi.org/10.3390/jpm10040203`.

7. Vlad Manea, Allan Berrocal, Katarzyna Wac. *Using Consumer-Friendly Wearables to Associate Patient- and Technology-Reported Physical Activity in Healthy Seniors*. Conference on Mobile Systems and Pervasive Computing (MobiSPC 2020). 8p. DOI: `https://doi.org/10.1016/j.procs.2020.07.036`.

8. Vlad Manea, Katarzyna Wac. *Using Consumer-Friendly Wearables to Associate Patient-Reported Quality of Life and Tech-Reported Physical Activity and Sleep in Healthy Seniors*. Poster at the Conference of the International Society for Quality of Life Research (ISOQOL 2020). DOI: `https://doi.org/10.1007/s11136-020-02626-y`.

9. Natalie Solomon, Vlad Manea. *Energy and Fatigue: Classification and Assessment of Energy and Fatigue using Subjective, Objective, and Mixed Methods towards Health and Quality of Life* (accepted). Book chapter in: Katarzyna Wac, Sharon Wulfovich (eds.), *Quantifying Quality of Life: Incorporating Daily Life into Medicine*, Health Informatics, Springer, Cham. 30p.

List of Figures

2 Motivation and Facilitation of Human Subject Study Participation

2.1 Background

Human subject studies are essential because repeated and harmful daily life behaviours may lead to disease in the long term [3]. The research on the motivation and facilitation of human subject study participation had the following objectives:

1. Review the literature on the factors, opportunities, and challenges of participation in human subject studies.

2. Propose model and framework designs that can improve human subject study participation by leveraging mobile and wearable technologies.

3. Extend the mQoL-Lab platform with tools that facilitate or support human subject study participation.

The research consisted of activities in two areas: the *motivation* and the *facilitation* of human subject study participation. In the area of *motivation* to participate in human subject studies, we assessed the willingness and motivation factors to participate in health studies and the challenges and opportunities of conducting longitudinal studies. We proposed a presentational model to improve the motivation to participate in health studies. Furthermore, we organised an international scientific workshop on longitudinal data collection in human subject studies, where we gathered a group of scientific experts to discuss experiences in this area.

Publications 1, 2, and 5 were part of the H2020 Wellbeing and Health Virtual Coach (WellCo, No. 769765, [16, 17]) research project. WellCo aimed to provide a novel ICT-based platform for wellbeing and health-oriented virtual coach for behaviour change. The publications used different designs; Publication 1 presented the mQoL-Chat chatbot design, while Publications 2 and 5 proposed the mQoL mobile app

design. Both artefacts were candidates for a study to co-calibrate PROs and TechROs collected as part of the WellCo project (Chapter 3 describes the co-calibration).

Publication 3 presents the Workshop on Longitudinal Data Collection in Human Health Studies (LDC 2019 [18]) organised by the QoL Lab and collaborators at the International Joint Conference on Pervasive and Ubiquitous Computing (ACM UbiComp 2019, [19]). We organised the workshop "*to bring together researchers involved in longitudinal studies to foster an insightful exchange of ideas, experiences, and discoveries, and discuss designs that may improve future studies' reliability, validity, and perceived meaning*" [20].

In the area of *facilitation*, we researched and developed the mQoL-Lab, a living lab platform embracing the factors for participation in human subject studies, and enabling to conduct the studies themselves.

Publication 4 describes the mQoL-Lab platform and living lab studies (from 2010 to 2020) leveraging mQoL-Lab. The motivation factors to participate in human subject studies informed the mQoL-Lab platform and its studies. In the publication, we shared the acquired experience from over ten studies conducted on the mQoL-Lab platform via "*requirements, architecture, design, step-by-step support, configuration notes, and recommendations for researchers to construct a software platform supporting human subject studies*" [21].

2.2 Methods

In **Publication 1**, we conducted a scoping literature review on the willingness and motivation to participate in human *health* studies from two perspectives: *populations* (healthy and diseased) and *object of data sharing* (electronic health records and wearable health data). We then proposed a presentational model based on personalised stories to improve retention and engagement in health study participants. Instead of only presenting the participants with the necessary steps in the study (i.e., a *uniform* experience across participants), the stories and moments motivate and incorporate them based on the their state (*personalised* experience). We provided three vignettes to compare the uniform and personalised experiences in the mQoL-Chat chatbot design for three *personas* [22]. We argued how mQoL-Chat could be a framework of choice for conducting health studies. Then, we discussed the advantages, challenges, and future avenues for research.

In **Publication 2**, we conducted a scoping literature review on *longitudinal health* studies the challenges and opportunities to assess individuals' QoL, informed by behaviours and health state measured in situ. We reviewed the opportunities and challenges from the perspectives of *motivation to participate* (leading to potentially high-quality and longitudinal data) and the *study's length* (expected to be positively affected by a high motivation). The paper then proposed the mQoL mobile app design as a literature review-informed unifying framework to conduct longitudinal studies addressing the challenges and leveraging the opportunities for participants and researchers. The same body of literature informed **Publication 5**.

In **Publication 3**, we described the *longitudinal* data collection workshop. The workshop paper reviewed the *human* and *technical* factors of participation in *longitudinal* studies and highlighted the challenges and opportunities in collecting longitudinal data. Then, it described the workshop objectives, contributions, and practicalities. Following the call for paper, we conducted two iterations of single-blind peer review for the submitted papers. Between the iterations, the authors revised the papers before the acceptance or rejection.

In **Publication 4**, we conducted a scoping literature review of the mobile sensing tools leveraged for human subject studies in situ. From the literature review and the challenges and opportunities encountered during over ten studies conducted on the mQoL-Lab platform, we derived and motivated requirements, the conceptual model, architectural design, and study design considerations of the mQoL-Lab platform. Then, the publication described in depth the technical implementation choices and critical learnings from the operationalisation of mQoL-Lab in the QoL Lab.

2.3 Results

Figure 2.1 overviews the relationships between the publications in this area and keywords describing the outcomes, models, and operationalisation of the proposed models and unifying framework. Appendix B.1 overviews all publications in this thesis. We derived *groups* defined as sets of two or more publications associated with the same set of keywords. Three groups emerged for this research area:

- **Group 1**: Publications reviewing the literature on the challenges and opportunities in conducting human subject studies (Publications 1, 2, 3).
- **Group 2**: Publications proposing designs that leverage the mQoL-Lab (Publications 1, 2, 4, 5).

- **Group 3**: Publications referring to a mobile app's design for conducting human subject studies leveraging and extending mQoL-Lab (Publications 2, 5).

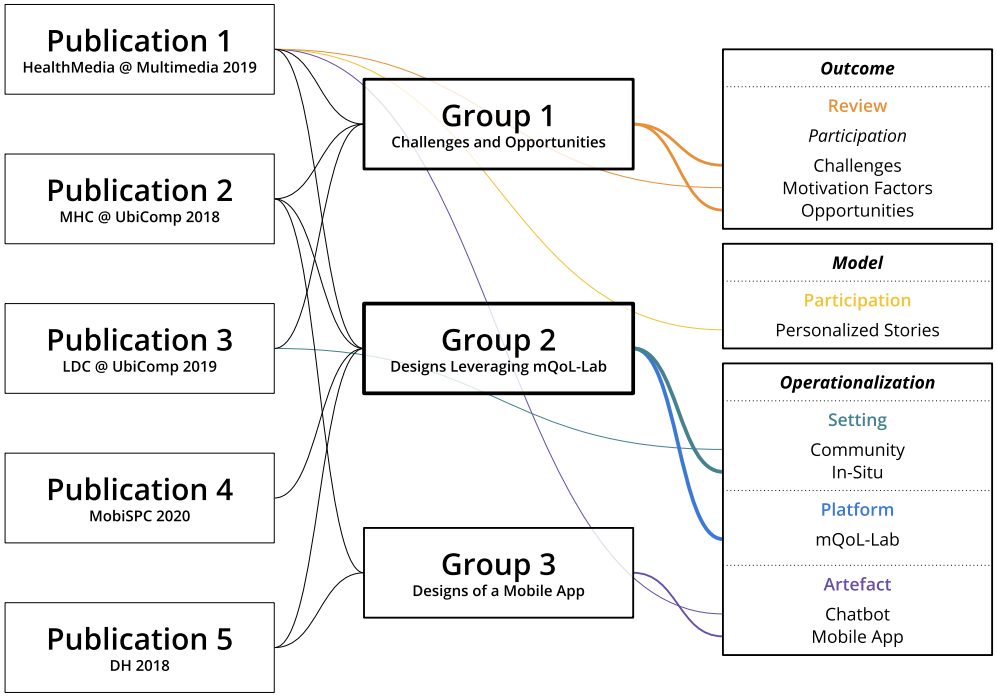
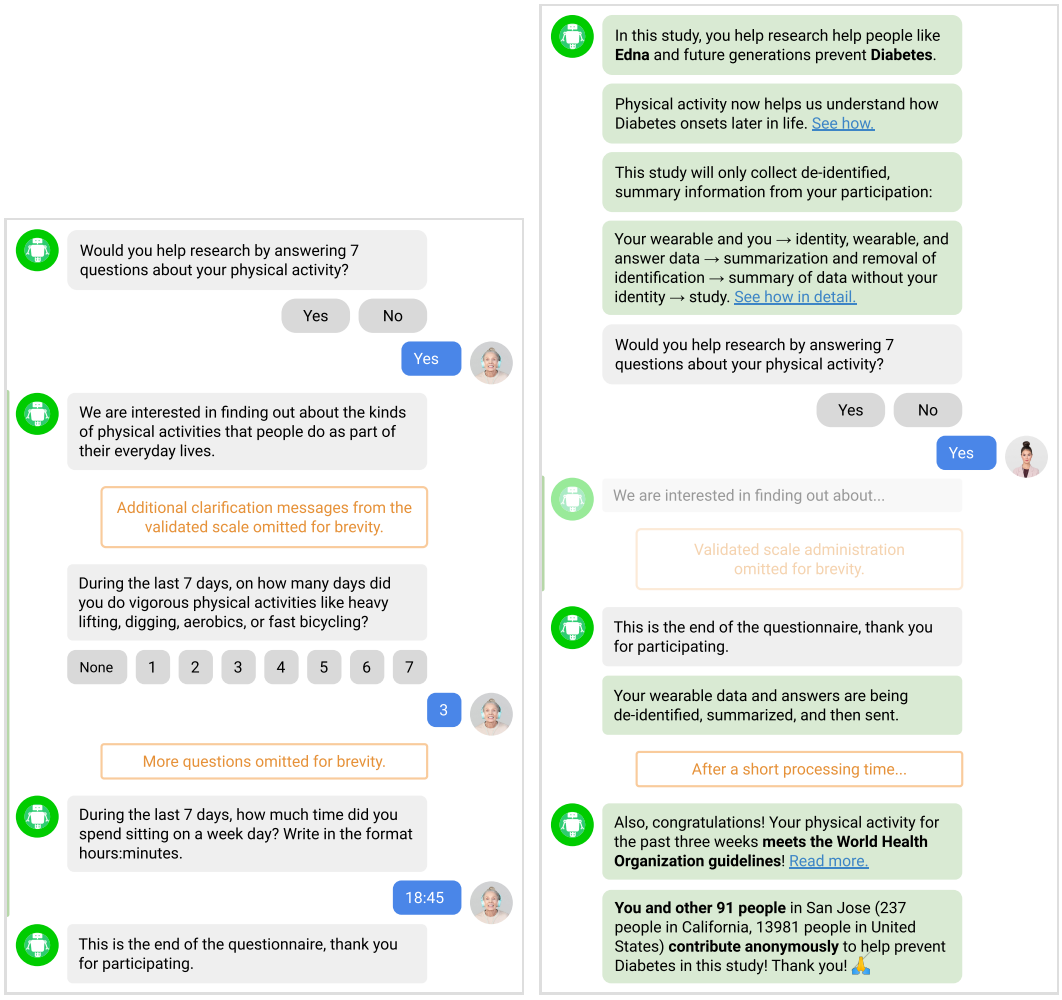


Fig. 2.1.: Publications (left), groups of publications (centre), and keywords (right) for the research on the motivation and facilitation of human subject study participation.

In **Publication 1**, we reviewed the willingness factors to share electronic health records and wearable health data, and the motivation factors to participate in *health* studies for healthy and diseased populations. The literature review found that participants are generally more willing to share their data with their primary health care providers than external entities. However, while they favour sharing their *complete* data with their primary care providers, they agree to share *parts* of their data with researchers, given a purpose and control of the data sharing. Altruistic motivation and personal non-financial benefits (qualitatively) overshadowed health gains and financial incentives. Participants are motivated by helping others and then, over time, by their benefits from the studies. Overall, the following four (qualitative) groups of motivation factors emerged from the literature review: *helping others* (e.g., family, friends, acquaintances, society), *personal benefits* (e.g., health, information, rewards, social), *study topic* (organ, process, behaviour, focus on QoL facets), and *study design and operationalisation* (e.g., delivery, effort, device, data). The paper then illustrated participation scenarios: uniform for all participants

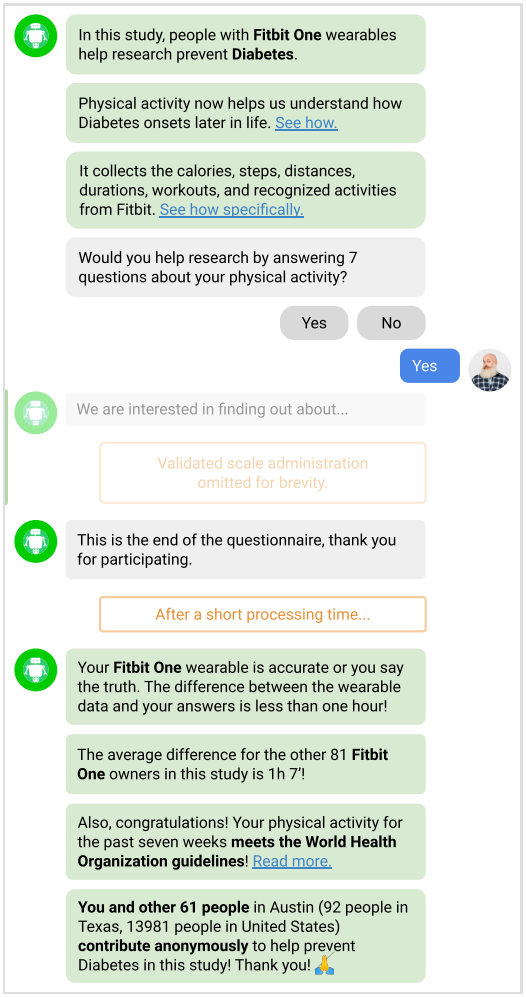
and then personalised based on three personas' motivation factors as depicted in Figure 2.2 in mQoL-Chat as illustrated by Figure 2.3.



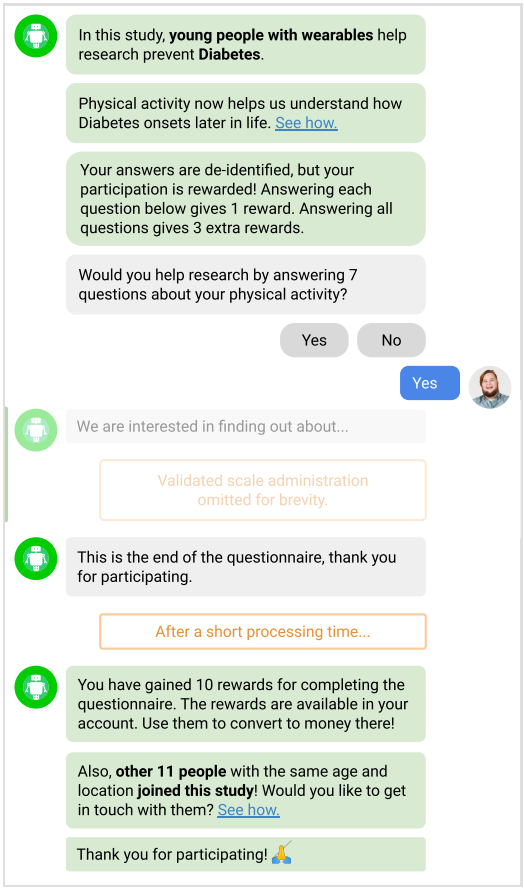
(a) Uniform story. While concise, this interaction does not personalise participation.

(b) Personalised story for Alice, a person with diabetes risk in her family [23].

Fig. 2.2.: Uniform story for all participants vs personalised stories for the Alice, Bob, and Charlie personas. The pale green line on the left side indicates a PRO's administration, characterised by little flexibility in the presented information. While less concise, the last three interactions personalise the participation with contextual moments (green). Extracted from Publication 1.



(c) Personalised story for Bob, a Quantified Self tracker [24] experimenting with wearables.



(d) Personalised story for Charlie, a student participating in studies for financial purposes.

Fig. 2.2.: Uniform story for all participants vs personalised stories for the Alice, Bob, and Charlie personas (continued). The pale green line on the left side indicates a PRO's administration, characterised by little flexibility in the presented information. While less concise, the last three interactions personalise the participation with contextual moments (green). Extracted from Publication 1.

In **Publication 2**, we reviewed the mobile health challenges faced by numerous app-based human subject studies. Highlighted challenges include the lack of scientific rigour, the lack of holistic assessment, and the burden on participants to manage an app for each study. Furthermore, data quality challenges included its long duration, high dimensionality, and levels of sensitivity. The paper then presented the mQoL mobile application and its architecture – see Figure 2.4 – to unify the requirements to conduct longitudinal health studies, including study management, collection of PROs, and collection of TechROs – see Figures 2.5 and A.1. The paper highlighted the opportunities for researchers and participants offered by

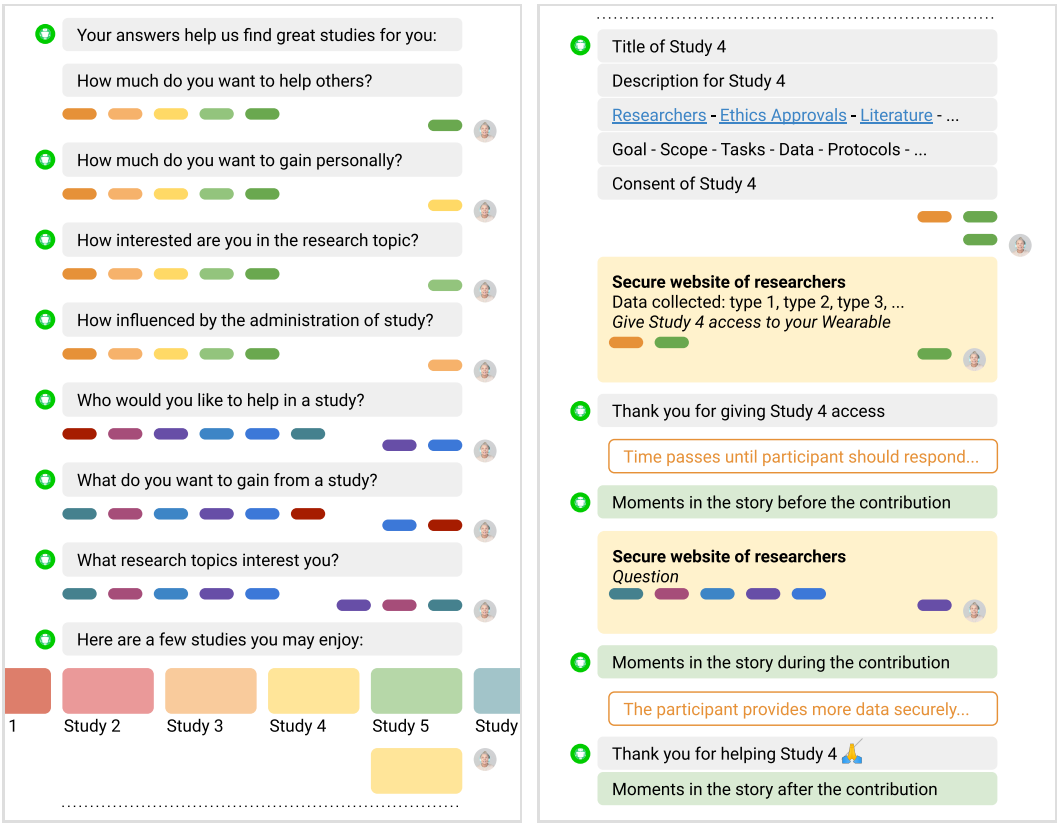


Fig. 2.3.: Chatbot unified human subject study format: identifying motivational factors for participation followed by study title, description, researchers, ethics, evidence, goal, scope, tasks, data, protocols, and consent. Extracted from Publication 1.

mQoL. Specifically, researchers can focus on the study instead of the platform, collect consented and pseudonymised data, and obtain longitudinal datasets in situ; participants can monitor, observe, and reflect upon their daily QoL using only evidence-based personalised information. **Publication 5** used a subset of these findings.

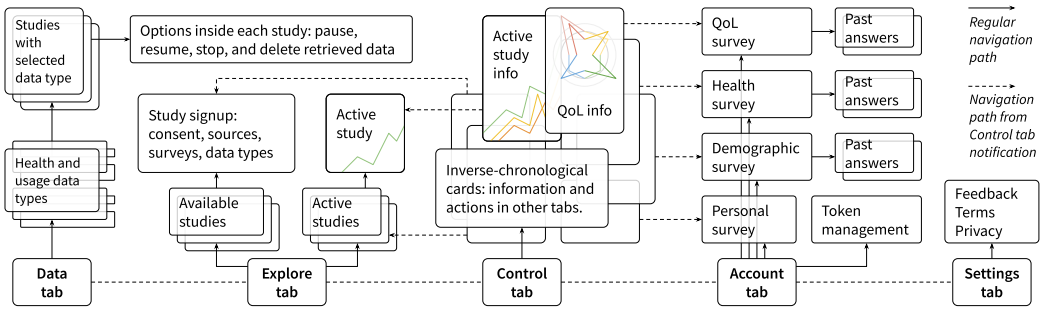


Fig. 2.4.: Conceptual architecture of the mQoL mobile app: tabs (*data*, *explore*, *control*, *account*, and *settings*) and modules. Extracted from Publication 2.

In **Publication 3**, we reviewed human factors and technical factors for *longitudinal* data collection in human subject studies. Human factors include the participants' attitudes towards self-monitoring, overall goals toward health and fitness, desires at a given time, and concerns over data security. Technical factors refer to both the data collection artefact and the properties of the data collected. Artefact-related factors include accuracy, usability issues, synchronisation problems, and battery life. Data-specific factors include data "*dimensionality, heterogeneity, temporal dependency, sparsity, irregularity, noisiness, ambiguity, and redundancy*" [25]. The International Workshop on Longitudinal Data Collection from Human Subject Studies "*foregrounded contributions and facilitated discussions focused on the methods, tools, and frameworks for collecting, analysing, and interpreting human subjects' data obtained over long periods*" [20] from three accepted papers. The workshop's agenda contained a keynote speech from a senior researcher specialised in longitudinal studies, presentations by the authors of three accepted papers, the best paper award, and a moonshot challenge where the participants designed a longitudinal study for the next 10-20 years. The workshop concluded with a social dinner.

In **Publication 4**, we derived requirements from three constituents: *participants*, *researchers*, and *system*, as seen in Figure 2.6. Participant requirements include consent, control of participation, control of data provision, and contribution to studies. Study researcher requirements included reusability, in-study data unification by participant, offline data collection, and adherence to data safety requirements. Finally, the system's requirements included managing different functionalities in studies with minimal programmatic intervention, the administration of interventions, and the

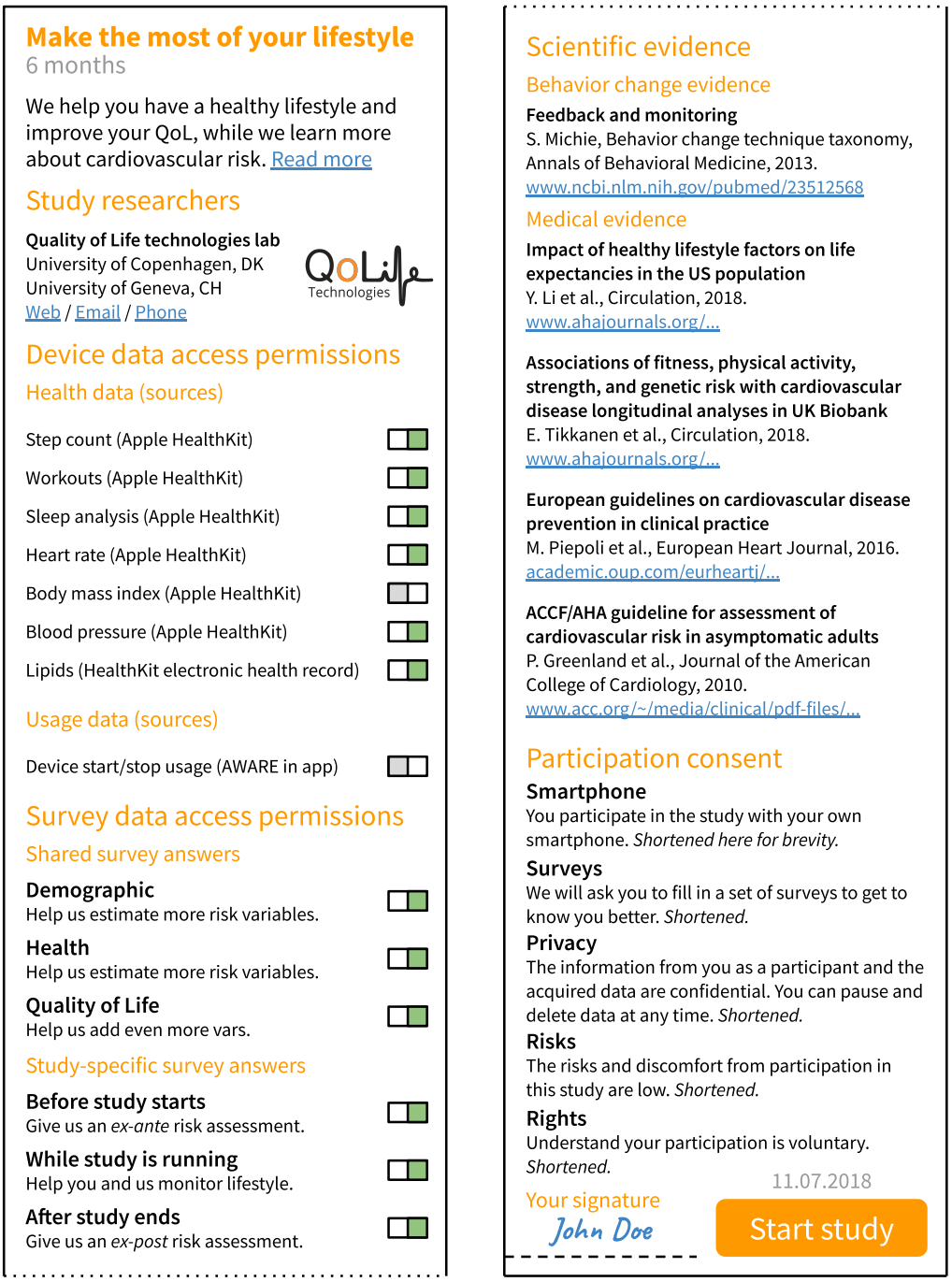


Fig. 2.5.: mQoL unified human subject study format: title, duration, description, researchers, permissions for device-reported (health and usage, TechRO) and self-reported (shared and study-specific, PRO) data, scientific evidence (behaviour change and medical), participation consent, and participant signature. Extracted from Publication 2.

ability to report data for analysis. The conceptual model of the architectural design consisted of *components* interconnected within two levels: *robust* and *flexible*. The robust layer serves as the foundation of the architecture. In each component, a set of transient *features* forms the flexible layer can be connected and disconnected upon changing research needs, as depicted in Figure 2.7. The component- and layer-based architectural design allows observational and interventional study designs, located in the lab and in situ, and leveraging PROs and TechROs collected continuously, scheduled, or event-based. mQoL-Lab served as the platform for over ten studies in four chronological stages: (1) single-study platform leveraging smartphone sensor-based data loggers, (2) platform enhancement with parallel studies with separate configurations, (3) platform enhancement with studies collecting consumer wearable data, and (4) platform re-instantiation in multiple geographical locations beyond the original location of Geneva, Switzerland.

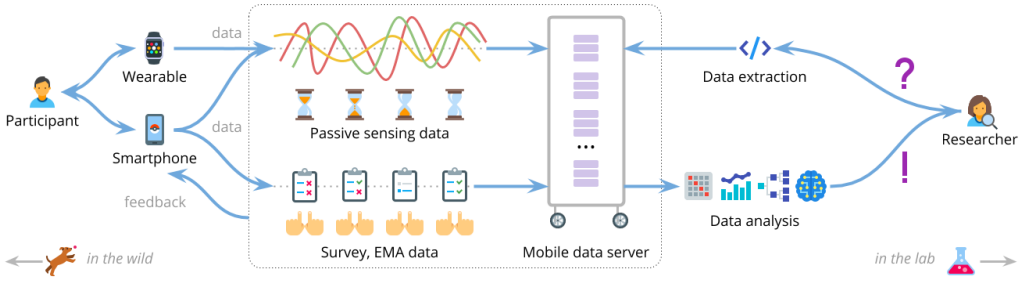


Fig. 2.6.: Overview of a human study conducted on the mQoL-Lab platform with the three constituents (participant, researcher, system) and provided data: passive sensing data (TechRO), and survey (PRO). Extracted from Publication 4.

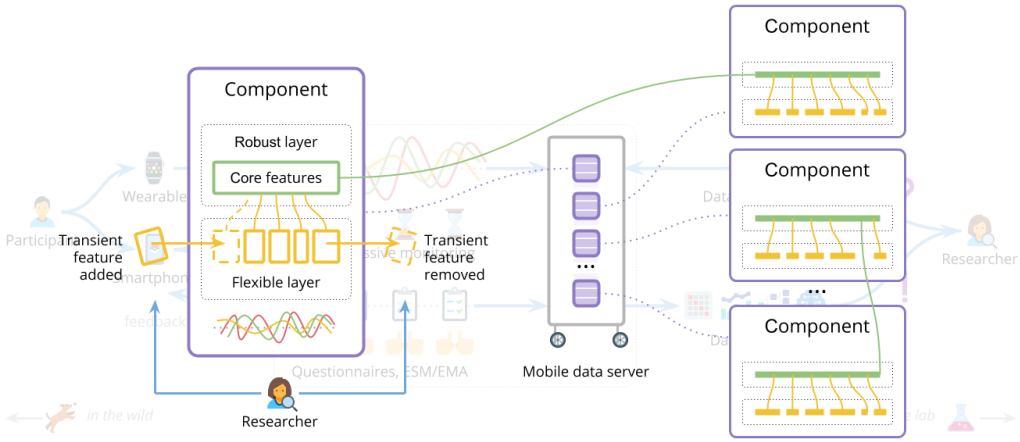


Fig. 2.7.: Architectural design of the mQoL-Lab platform. The components are represented in magenta. The core features in the robust layer are depicted in green. The transient features in the flexible layer are shown in yellow. Extracted from Publication 4.

In addition to the published work, the research activities included **two master theses** at the Departments of Communication and Computer Science, University of Copenhagen.

1. In a master thesis entitled *Designing for Participation in Longitudinal Health and Wellbeing Studies* [26], Alba Kejser Perez, Cecilie Rosentoft, and Elisabeth Brinth Refstrup studied the factors of motivation for participation in health and wellbeing studies by employing qualitative and quantitative methods (*mixed methods*). This thesis used Publication 1 as a basis for research.

2. In a master thesis entitled *Engaging Participants in the Recruitment Phase of Human Subject Health Studies – mQoL-Chat: a Chatbot Approach* [27], Mads Schnoor Hansen implemented a minimum viable version of the chatbot that collected PRO and TechRO data, by leveraging the two designed and developed modules in the mQoL-Lab (Chapter 4). This thesis used Publications 2 and 5 as a basis for research. Mads then contributed to Publication 1.

2.4 Discussion

Publications 1, 2, and 5 used a chatbot and a mobile app as unifying framework artefact designs for conducting human subject studies. Conducting studies using implementations of these designs provides numerous advantages for study participants, researchers, and developers. For both designs, participants can benefit from easy access and timely support. For the mobile app (the preferred choice in 2018), they benefit from the app's familiarity on the smartphone. For the chatbot (the preferred choice in 2019), they benefit from communication in natural language, and a simple user interface, based on chat messages and limited media. Researchers benefit from the flexibility and expressiveness of the conducted studies, and in the case of the chatbot, access to its underlying social network. For the mobile app, developers can benefit from the numerous existing data collection platforms for both PROs and TechROs (e.g., ResearchKit [28] and HealthKit [29] for the iOS platform). Using the chatbot, developers can prototype, experiment, and implement presentations and interactions with ease, using either a wizard or code (e.g., Chatbot [30]).

The personalised stories presentational model in the mQoL-Chat unifying framework proposed in **Publication 1** poses several challenges. First, personalised stories and moments require significantly more content than a study administration uniform for all participants. Then, the personalised stories presented in the chatbot design may introduce several selection biases in the participants (due to both human

and technological factors). Finally, the data collection using a chatbot requires a transparent data management procedure from the chatbot user interface to the mQoL-Lab platform and back.

The empirical study of participation in human subject studies by using the mobile app or chatbot design with personalised stories, by balancing the personalisation of the study for the participants with the operationalisation burden for the researchers and developers, is an avenue for future research.

Long-term, researchers can use properties of the *participant* (e.g., demographics, topics of interest, factors, and concerns for participation, reasons for abandonment), initial features of *study* (e.g., goals, story, requirements, rewards), and properties of *involvement* (e.g., responses to moments as markers for engagement and retention, the outcome of completion or attrition) that change in time. Supervised learning models that predict the likelihood to participate in a study for a given duration and then recommend the fittest studies for future participation are another direction for future research subject to the underlying data's availability.

The workshop activity described in **Publication 3** foregrounded state of the art contributions in collecting data over long periods. However, it occurred only one time. A future recurring event spanning ten years (mirroring longitudinal studies), with a periodicity of 3-12 months, would allow a cohort of 10-100 of research labs to report ongoing progress, share recent insights, and adjust selected aspects of their study designs to address old and new challenges and leverage new opportunities in human subject study participation.

Researchers with long-term goals in human subject studies will benefit from building a reliable and scalable architecture that supports their growing needs, such as mQoL-Lab. The mQoL-Lab platform described in **Publication 4** is relevant in particular for researchers preparing to conduct human subject studies involving simultaneous participants (tens to hundreds), from a few days to several years, by employing qualitative, quantitative, or mixed research methods, potentially across geographies.

We plan to extend the scientific contribution described in Publication 4 on mQoL-Lab with videos, tutorials, and code snippets that can further empower this research community.

The two master theses contain significant findings on the motivation of human subject study participation using mixed methods (N = 100) not included in this PhD thesis. We plan to write a joint manuscript on the motivation for participation in human subject studies based on their findings.

Following the Research App by Apple in October 2019 [31], Google announced the new Google Health Studies mobile app in December 2020 [32]. The intent and design of both applications resemble the mQoL design proposed in Publications 2 and 5. Table B.1 highlights similar paragraphs from the descriptions of the mQoL design [33], Apple app [31], and Google app [34]. These releases by two large companies of mobile apps bearing similarities with mQoL (unifying frameworks for human subject studies) confirm the relevance and feasibility of the mQoL design.

2.5 Scientific Contributions

We contributed to the state of the art in studying the motivation and facilitation of human subject study participation through five publications, as depicted in Figure 2.8. Appendix B.2 overviews all publications and scientific contributions in this thesis.

Two groups of publications providing the same scientific contributions emerged:

- **Group 1**: Publications thoroughly exploring the research area of motivation for participation in human subject studies (Publications 1, 2, 3).
- **Group 2**: Publications proposing a mobile application design as a unifying framework for longitudinal human subject studies (Publications 2 and 5).

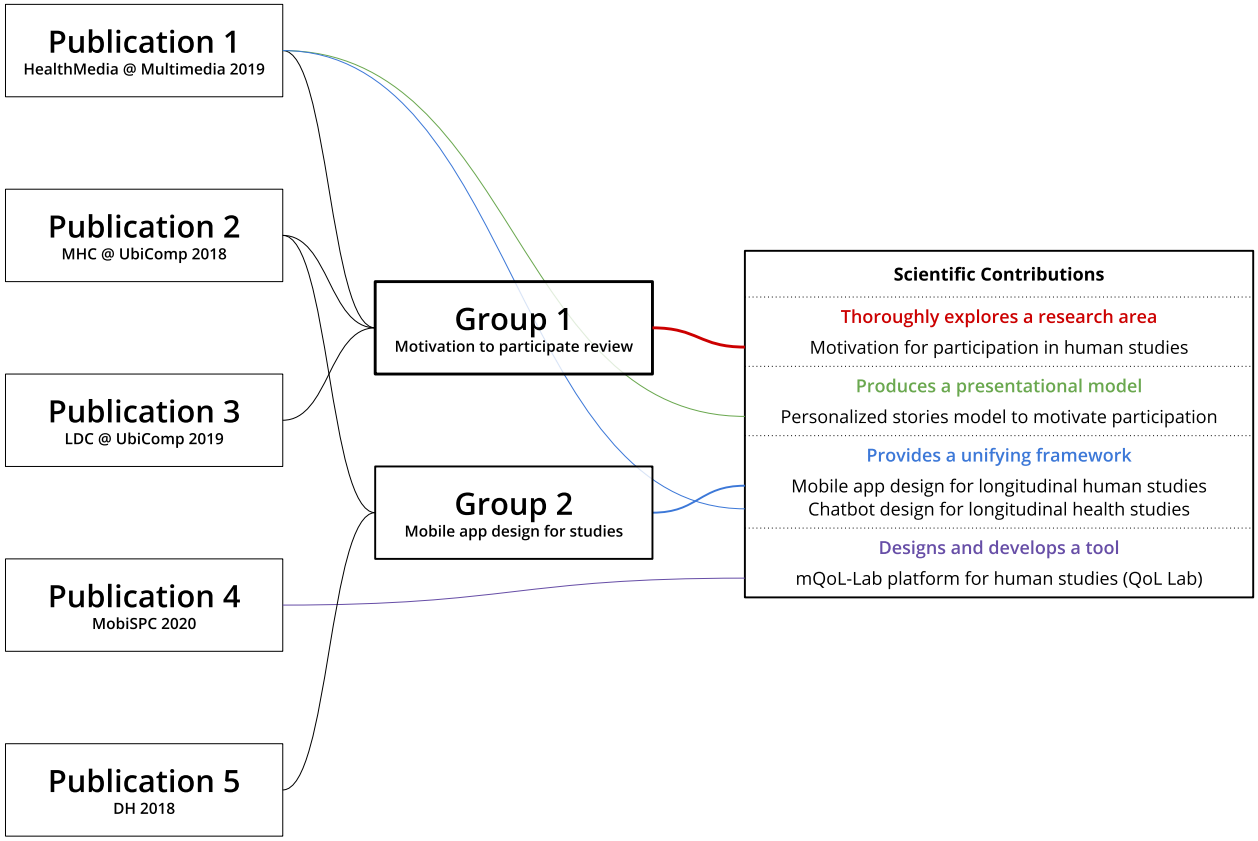


Fig. 2.8.: Publications (left), groups of publications (centre), and scientific contributions (right) for the research on the motivation and facilitation of human subject study participation.

The publications, along with their scientific contributions to state of the art on the motivation and facilitation of human subject study participation, were:

1. Vlad Manea, Mads Schnoor Hansen, Semahat Ece Elbeyi, Katarzyna Wac. *Towards Personalizing Participation in Health Studies*. Workshop on Multimedia for Personal Health and Health Care (HealthMedia 2019), Conference on Multimedia (MM 2019). 8p. DOI: `https://doi.org/10.1145/3347444.3356241`. This publication can be found in Appendix A. [35]

    - Thoroughly explores the research area of motivation for participation in human subject studies (Group 1).
    - Produces a presentation model using personalised stories to motivate participation in health studies.
    - Provides the mQoL-Chat chatbot design as a unifying framework to conduct longitudinal health studies.

2. Vlad Manea, Katarzyna Wac. *mQoL: Mobile Quality of Life lab: from Behavior Change to Quality of Life*. Workshop on Mobile Human Contributions (MHC 2018), Conference on Pervasive and Ubiquitous Computing (UbiComp 2018). 6p. DOI: `https://doi.org/10.1145/3267305.3267549`. This publication can be found in Appendix A. [33]

    - Thoroughly explores the research area of motivation for participation in human subject studies (Group 1).
    - Proposes a mobile application design as a unifying framework for longitudinal human subject studies (Group 2).

3. Vlad Manea, Allan Berrocal, Alexandre De Masi, Naja Holten Møller, Katarzyna Wac, Hannah Bayer, Sune Lehmann, Euan Ashley. *Call for Papers: LDC '19: Workshop on Longitudinal Data Collection in Human Subject Studies*. Call for Papers for the Workshop on Longitudinal Data Collection in Human Subject Studies (LDC 2019), Conference on Pervasive and Ubiquitous Computing (UbiComp 2019). 4p. DOI: `https://doi.org/10.1145/3341162.3347758`. This publication can be found in Appendix A. [20]

    - Thoroughly explores the research area of motivation for participation in human subject studies (Group 1).

4. Allan Berrocal, Vlad Manea, Alexandre De Masi, Katarzyna Wac. *mQoL-Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices*. Conference on Mobile Systems and Pervasive Computing (MobiSPC 2020). 9p. DOI: `https://doi.org/10.1016/j.procs.2020.07.033`. **Nominated for the best project pitch award at the University Hospitals of Geneva Innovation Day 2020, Geneva, Switzerland**. This publication can be found in Appendix A. [21]

    - Designs and develops the mQoL-Lab platform as a tool for conducting human subject studies.

5. Vlad Manea, Vero Estrada-Galiñanes, Katarzyna Wac. *mQoL: Mobile Quality of Life lab*. Poster and demo at the Digital Health Conference (DH 2018). **Nominated for the Innovation Prize in the category of the best data-driven innovation, Lyon, France**. This publication can be found in Appendix A.

    - Proposes a mobile application design as a unifying framework for longitudinal human subject studies (Group 2).

3 Co-Calibration of Behavioural, Health, and Quality of Life Outcomes

3.1 Background

Little research focused on assessing the relationships between sets of different behavioural, health, and QoL outcomes assessed via PROs and consumer wearable TechROs in healthy seniors, in situ and longitudinally.

The research met the following objectives:

1. Review the literature on the co-calibration of behavioural, health, and QoL outcomes in seniors.
2. Provide a computational model to co-calibrate small samples of PROs and TechROs obtained from validated scales and consumer wearables, respectively.
3. Demonstrate the computational model's feasibility on the dataset resulting from an observational human subject study on healthy seniors, longitudinally, and in situ.
4. Collect the behavioural, health, and Quality of Life PROs by leveraging the mQoL-Lab platform.
5. Collect the digital biomarker TechROs by leveraging the mQoL-Lab platform.
6. Assess data quality properties of the PROs and TechROs collected in the study.
7. Assess the co-calibration between PROs and TechROs by using statistical methods appropriate for the data quality properties above.
8. Inform longitudinal and in situ studies' design by leveraging the data quality and co-calibration patterns from the previous two objectives.
9. Review in-depth the literature on assessing a QoL facet by using a combination of reported outcomes.

In **Publications 6, 7, and 8**, we researched as part of the EU AAL Caregiver and ME (CoME, No. 14-7, 2017–2020) research project. CoME aimed at the self-management of health for healthy seniors, but at risk of mild cognitive impairments, and their informal caregivers [36] to reduce the long-term risk of severe disease (e.g., dementia) and improving Quality of Life. The institutional review board at the University of Geneva approved the project in 2016. Participants were from Hungary and Spain. All individuals signed written consent before participating.

The project used numerous PROs to obtain a holistic view of the participants' behaviours, health, and QoL, by covering constructs that are *reflective* (e.g., physical activity, anxiety, depression, memory, sleep) or *formative* (e.g., nutrition, social support) [37, 38]. These constructs assess participants' health state and correspond to several behavioural risk factors of dementia [39], as guided by the project's goals.

In **Publication 9**, we conducted a scoping review on the technology-enabled assessment of the *Energy and Fatigue* facet in the Quality of Life model of the World Health Organisation [40]. This work serves as a prerequisite for the literature review on co-calibration as well as an individual contribution.

3.2 Methods

In **Publications 6, 7, and 8**, we conducted a longitudinal observational study in situ by leveraging mixed methods. We collected *qualitative* PROs (physical activity, social support, anxiety and depression, memory, nutrition, sleep, and health-related QoL) and *quantitative* TechROs (e.g., physical activity, sleep, and heart rate) for seniors, self-reported healthy or with mild disease. We assessed the PROs by using validated scales (8 questionnaires with validated scales: IPAQ for physical activity [41], MSPSS for social support [42], GADS for anxiety and depression [43], PREDIMED for Mediterranean nutrition [44, 45], SelfMNA for nutrition [46], MFE for memory [47], PSQI for sleep [48], EQ-5D-3L for health-related QoL [49]) and TechRO digital biomarkers (physical activity, sleep, heart rate) longitudinally by using consumer wearables (Fitbit Charge 2 [50]), respectively.

We aligned the PRO answers (for questions and their associated scores and subscores) with TechRO intervals of various durations (7-120 days, beyond the recall periods of the PROs) that ended on the administration date of the PRO (within a leeway). We included TechROs in both absolute amounts and relative amounts (e.g., the ratio of sedentary duration over the other durations up the 24 hours in the day

[51]) We applied descriptive statistics (e.g., mean, median, standard deviation) to assess the quality of the obtained data (PROs, TechROs, alignments of PROs and TechROs). Then, we applied inferential statistics through hypothesis testing (for this study, Spearman non-parametric rank correlations for at least ordinal variables corresponding to the assessed outcomes) to co-calibrate PROs with TechROs. We used the strong, significant correlations between PROs and TechROs to observe patterns of correlations (in this study, using two metrics: counting strong correlations and observing spectrums of correlations). We denote the aforementioned computational model as *coQoL*. Figure 3.1 illustrates a diagram of coQoL.

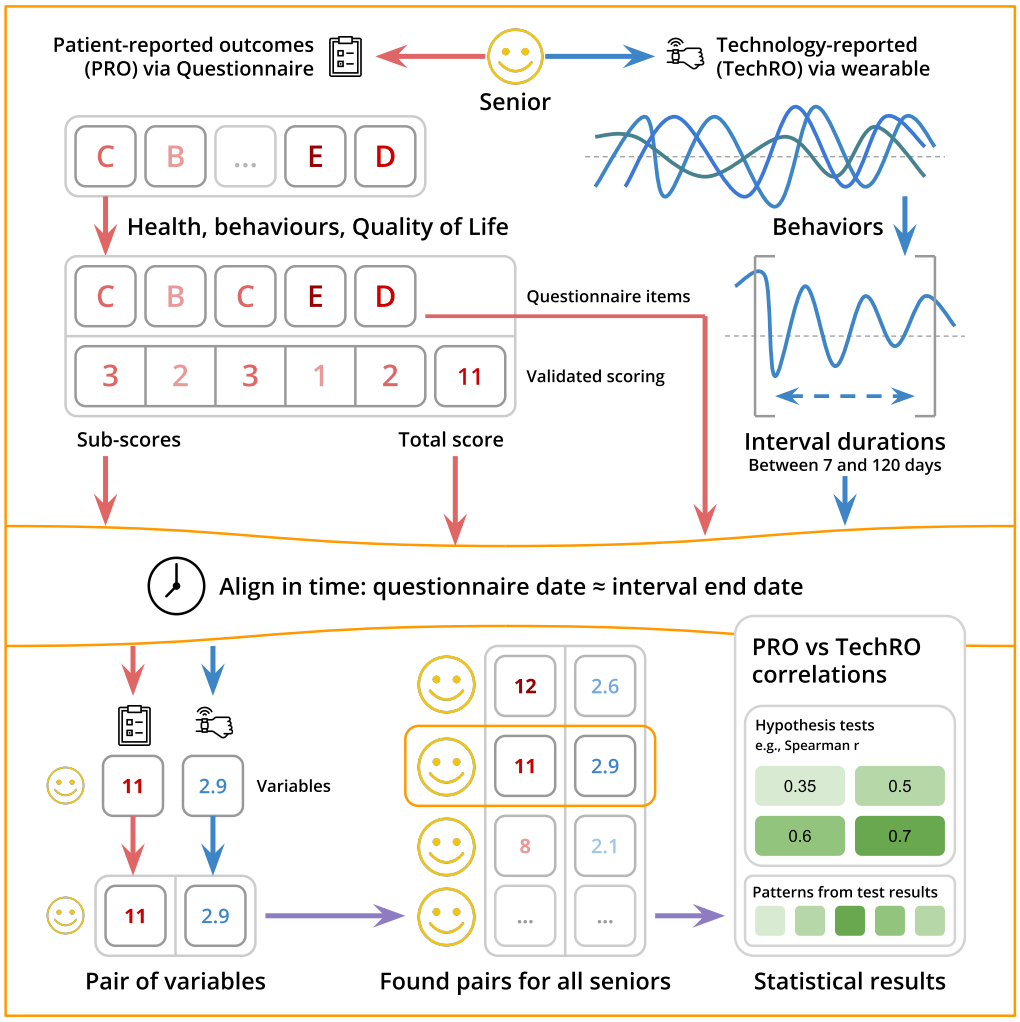


Fig. 3.1.: coQoL computational model for the co-calibration of PROs and TechROs.

In **Publication 9**, we conducted a scoping review of the recent literature (2010-2020) on the technology-enabled assessment of energy and fatigue, using domains (*keywords*) such as energy and fatigue (e.g., energy, fatigue, fatigability, tiredness,

vitality), populations (e.g., athlete, driver, performance, pilot, police, shift, sport, worker, employee), health outcomes (cancer, cardiovascular, circulation, dementia, heart, kidney, mental, pulmonary, respiration), and measurement (e.g., accelerometer, app, application, camera, band, ecological momentary assessment, performance, capacity, electrocardiogram, electrooculogram, experience sampling method, Fitbit, galvanic, mobile, sensor, smart band, smartphone, smartwatch, vision, watch, wearable).

3.3 Results

Figure 3.2 overviews the relationships between the publications in this area and keywords describing the proposed model's outcomes, models, and operationalisation. Appendix B.1 overviews all publications in this thesis. We derived three groups for this research area:

- **Group 4**: Publications assessing PRO of physical activity by using the IPAQ scale (Publications 6, 7).
- **Group 5**: Publications assessing PRO of health-related QoL by using the EQ-5D-3L scale and TechRO of heart rate by using the Fitbit Charge 2 consumer wearable (Publications 6, 8).
- **Group 6**: Publications assessing PROs and TechROs using the coQoL computational model (Publications 6, 7, 8).

In **Publication 6** (8 PROs), we reported that 39 seniors provided on average 7.4 $\pm$ 4.4 PROs for physical activity (IPAQ), social support (MSPSS), anxiety/depression (GADS), nutrition (PREDIMED, SelfMNA), memory (MFE), sleep (PSQI), Quality of Life (EQ-5D-3L), and 295 $\pm$ 238 days of TechROs (Fitbit Charge 2) along two years. We co-calibrated PROs and TechROs (coQoL) and reported human factors guiding coQoL use. We report high PRO—TechRO Spearman correlations ($r_S \geq 0.8$, $p < 0.05$) for physical activity (moderate domestic activity—light+fair active duration), social support (family help—fair activity), anxiety/depression (numeric score—sleep duration), or sleep (duration to sleep—sleep duration) at various durations (7–120 days).

In **Publication 7** (1 PRO: physical activity), we quantified the relations between physical activity outcomes, as patient-reported by 31 seniors (mean age 70.6 $\pm$ 3.2) through 53 answers (1.71 $\pm$ 0.96 / person) on the International Physical Activity Questionnaire (IPAQ) with a 7-day recall period, and 5615 days (mean 181.1 $\pm$

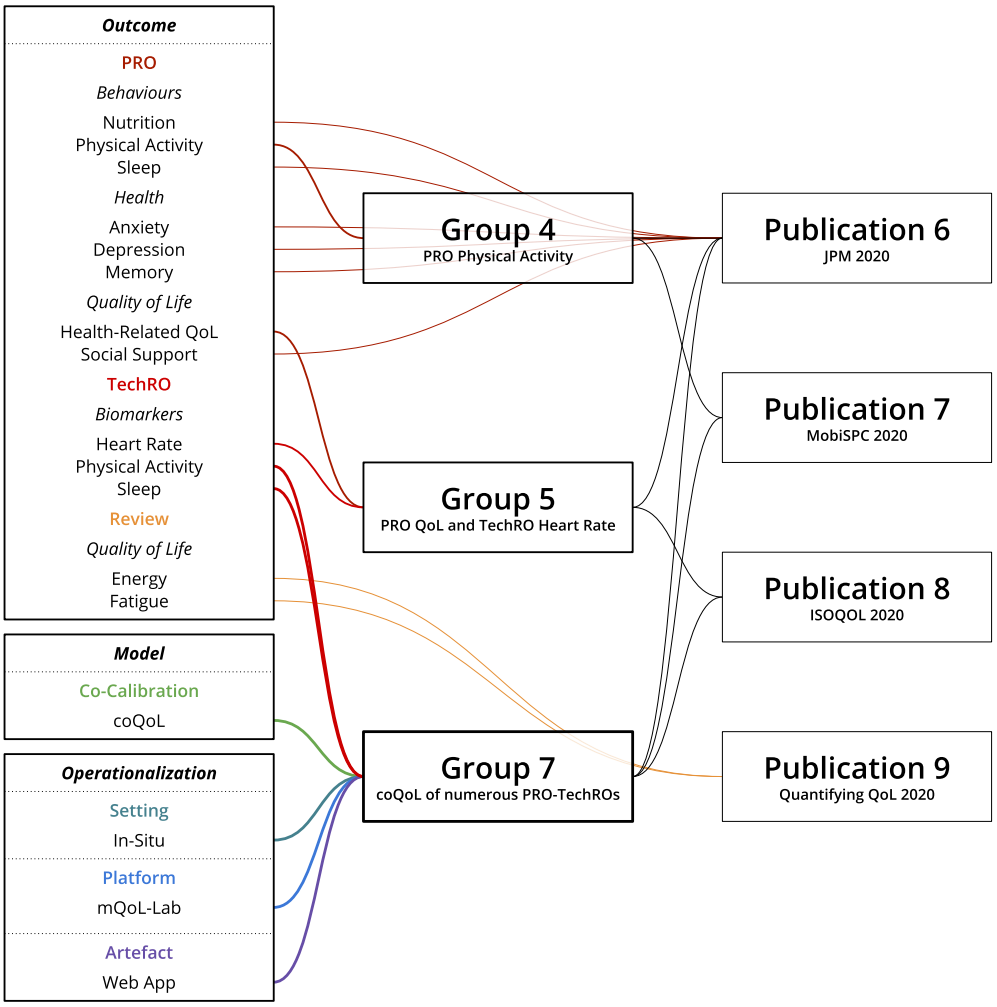


Fig. 3.2.: Publications (left), groups of publications (centre), and keywords (right) for the research on the co-calibration of behavioural, health, and Quality of Life outcomes.

179.2 days collected /person) technology-reported by Fitbit Charge 2. The wearables monitored daily life behaviours of physical activity and sleep for long durations (7 to 120 days). We applied coQoL. We found strong Spearman correlations between light and moderate IPAQ physical activity in the domestic activity domain, and light-fair intensity Fitbit physical activity (e.g., r_S = 0.88, p < 0.005). We also found negative moderate-strong correlations between Fitbit sedentary duration and all IPAQ physical activity domains and intensities (e.g., r_S = 0.64, p < 0.005).

In **Publication 8** (1 PRO: health-related Quality of Life), 31 seniors (mean age 70.66 ± 3.15) provided 54 EQ-5D-3L answers (1.72 ± 1.12 / person) and 9.150 Fitbit Charge 2 days (295.16 ± 247.25 / person). We applied coQoL. For the healthy participants, we found strong Spearman correlations (p < 0.05) for the PRO pain / discomfort vs the TechRO absolute sedentary duration (r_S = 0.69), mobility vs absolute sedentary duration (-0.57), and health state vs heart rate (r_S = -0.56). For the participants with self-reported mild disease, we found strong correlations (p < 0.05) for mobility vs steps (r_S = 0.71), distance (r_S = 0.71), and absolute sedentary duration (r_S = -0.67) as well as anxiety / depression vs steps (r_S = -0.57) and distance (r_S = -0.62). For all participants, pain / discomfort vs relative fair activity (r_S = 0.69) and sleep (r_S = -0.58); health state vs relative light activity duration (r_S = 0.63) and sleep duration (r_S = 0.73) yielded strong correlations (p < 0.05).

In **Publication 9**, we found 40 reviews on energy and fatigue and 60 studies assessing fatigue using technology. We classified fatigue as pathological and non-pathological, and then as physical and mental. We assessed the *qualitative* (subjective), *quantitative* (objective), and mixed-methods fatigue measurement methods (scale instruments, momentary assessments, physical assessment, cognitive assessment, cardiac physiology, ocular physiology, neural physiology, biologic markers, and behavioural markers). Finally, we placed the measurement methods on a spectrum based on several properties (validated, quantifiable, frequent, continuous, judgement-free, mood-free, memory-free, owned, and contextual).

In addition to the published work, the research on co-calibration formed the basis of **a master thesis** at the Department of Computer Science, University of Copenhagen. In the master thesis entitled *ConsistencyQoL – A Framework For Modelling Consistency In Behavioural Data Collected With Wearables* [52], Kirke Kjellberg designed a system for monitoring the consistency in human behaviours reported by wearables towards its co-calibration as a TechRO with PROs, by leveraging the two mQoL-Lab data collection modules (Chapter 4).

3.4 Discussion

In **Publications 6, 7, and 8**, we demonstrated the feasibility of coQoL. The computational model could identify numerous relationships between PROs and TechROs in healthy seniors longitudinally and in situ. Furthermore, we reported the data quality and potential human factors (primarily in Publication 6). Finally, we provided patterns of relationships that can inform future observational (and, where relevant, interventional) study designs for healthy seniors to assess and improve health and QoL.

coQoL is a robust computational model due to the flexibility of the co-calibration process. We illustrate in Figure 3.3 several parameters with which the range of human subject studies using PROs and TechROs that can leverage coQoL can expand beyond the study presented in this thesis.

A series of limitations characterised the study: the small sample size, the reduced power, and the apparent simplicity of the underlying methods for co-calibration (descriptive and inferential statistics). We addressed these limitations by analysing on numerous increasing duration intervals and decreasing number of observations, and allowing leeways to align PRO answers and TechRO intervals. The study highlights the challenge of retaining individuals, shared by numerous human subject studies, and motivating the research's relevance in Chapter 2.

While past studies on seniors may have had larger sample sizes than those in this study, they have not yielded stronger statistical results. For example, other co-calibration studies rarely report Spearman values of $r_S > 0.5$. Also, past studies assessed the behaviours for reduced durations (7-14 days). However, the study duration of over a few weeks is essential to overcome the "novelty" effect of the technology (TechRO) on the state and behaviour of the senior [53]. Conversely, the research co-calibrated (PRO) behaviour and health outcomes in healthy seniors longitudinally (up to 120 days) and in situ (while daily life unfolded). Nevertheless, the study focused on patterns of relations and not individual relations between PROs and TechROs.

In **Publication 9**, we did not observe validated calibration between objective measures and the concept of energy and fatigue. Therefore, both qualitative and quantitative measurements are being considered in the context of energy and fatigue, as they are essential, and so far semantically separate, indicators for health and QoL.

In an ongoing research project, we are using coQoL to assess relationships between heart disease events and behaviours using consumer wearables. The project is a

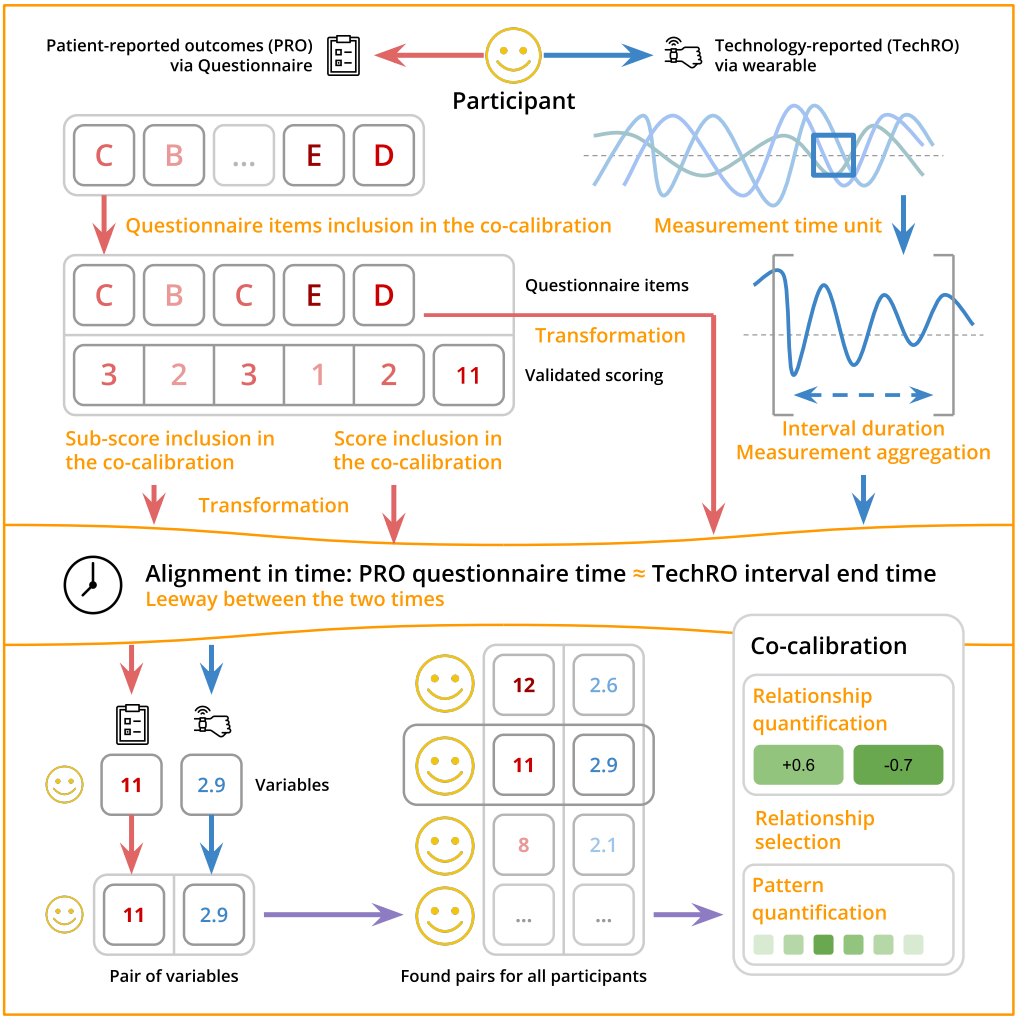


Fig. 3.3.: Parameters for the coQoL computational model (orange) for the co-calibration of PROs and TechROs. This illustration is not peer-reviewed.

collaboration between the University of Copenhagen and the Vital Beats private company [54], where the author of this PhD thesis had the mandatory research stay in an environment external to the University of Copenhagen.

We expect to design future studies with more participants for shorter periods (60–90 days), repeated every few months to a year, and focus on the PROs and TechROs delineated by the patterns reported by this research. Larger sample sizes will allow more advanced techniques. Finally, we aim to derive co-calibration-based trajectory models for individuals and populations.

3.5 Scientific Contributions

We contributed to state of the art in the study of co-calibration of behavioural, health, and QoL outcomes through four publications, as depicted in Figure 3.4. Appendix B.2 overviews all publications and scientific contributions in this thesis.

One group of publications providing the same scientific contributions emerged:

- **Group 3**: Publications producing the coQoL computational model for PRO-TechRO co-calibration, providing empirical data as patterns of PRO-TechRO statistical correlation, and leveraging the two mQoL-Lab modules as tools for questionnaire and wearable data collection (Publications 6, 7, 8).

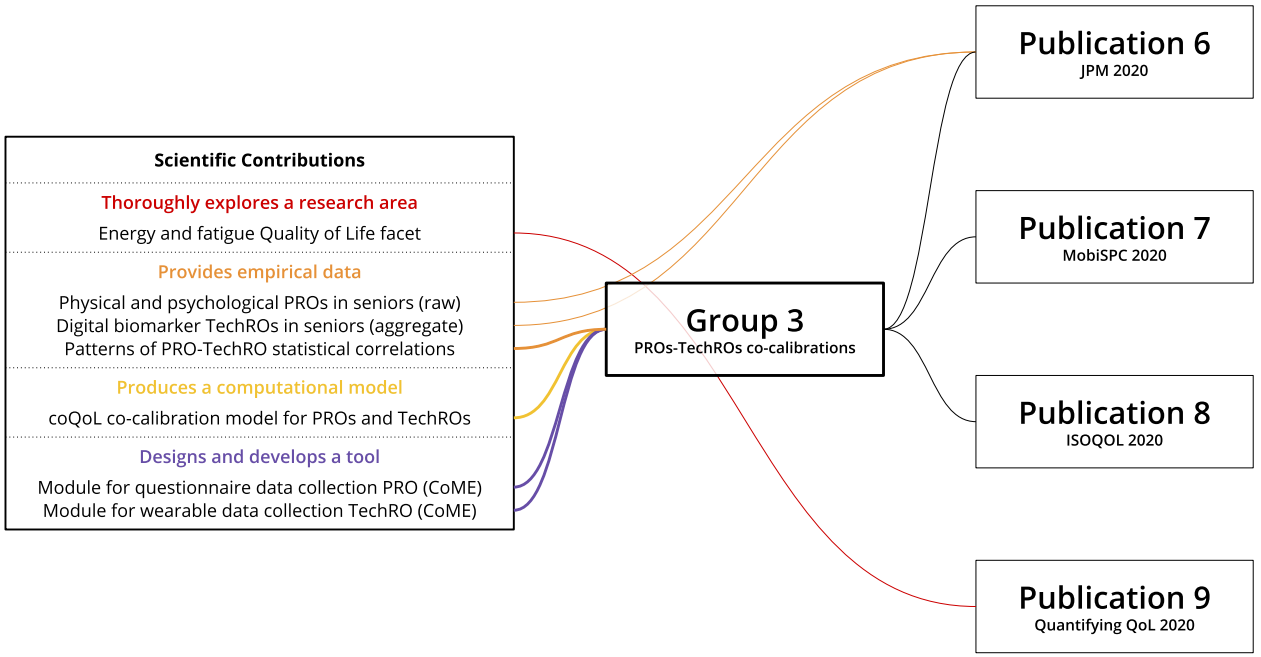


Fig. 3.4.: Publications (left), groups of publications (centre), and scientific contributions (right) for the research on the co-calibration of behavioural, health, and Quality of Life outcomes.

The publications, along with their scientific contributions to state of the art on the co-calibration of behavioural, health, and QoL outcomes, were:

6. Vlad Manea, Katarzyna Wac. *Co-calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method*. Journal of Personalized Medicine, 10(4), MDPI, 2020. Special Issue: *PROomics: Patient Reported Outcome (PRO) and Self-Tracking for Personalized Medicine*. Impact factor 4.433, rank 10/102 (Q1) in Health Care Sciences and Services. 41p. DOI: `https://doi.org/10.3390/jpm10040203`. This publication can be found in Appendix A. [55]

    - Produces the coQoL computational model for PRO-TechRO co-calibration (Group 3).
    - Provides empirical data as patterns of PRO-TechRO statistical correlation (Group 3).
    - Leverages the two mQoL-Lab modules as tools for questionnaire and wearable data collection (Group 3).
    - Provides empirical data as physical and psychological qualitative PROs (raw) in seniors, collected longitudinally in situ.
    - Provides empirical data as digital biomarker quantitative TechROs (aggregate) in seniors, collected longitudinally in situ.

7. Vlad Manea, Allan Berrocal, Katarzyna Wac. *Using Consumer-Friendly Wearables to Associate Patient- and Technology-Reported Physical Activity in Healthy Seniors*. Conference on Mobile Systems and Pervasive Computing (MobiSPC 2020). 8p. DOI: `https://doi.org/10.1016/j.procs.2020.07.036`. This publication can be found in Appendix A. [56]

    - Produces the coQoL computational model for PRO-TechRO co-calibration (Group 3).
    - Provides empirical data as patterns of PRO-TechRO statistical correlation (Group 3).
    - Leverages the two mQoL-Lab modules as tools for questionnaire and wearable data collection (Group 3).

8. Vlad Manea, Katarzyna Wac. *Using Consumer-Friendly Wearables to Associate Patient-Reported Quality of Life and Tech-Reported Physical Activity and Sleep in Healthy Seniors*. Poster at the Conference of the International Society for Quality of Life Research (ISOQOL 2020). DOI: `https://doi.org/10.1007/s11136-020-02626-y`. This publication can be found in Appendix A. [57]

    - Produces the coQoL computational model for PRO-TechRO co-calibration (Group 3).
    - Provides empirical data as patterns of PRO-TechRO statistical correlation (Group 3).
    - Leverages the two mQoL-Lab modules as tools for questionnaire and wearable data collection (Group 3).

9. Natalie Solomon, Vlad Manea. *Energy and Fatigue: Classification and Assessment of Energy and Fatigue using Subjective, Objective, and Mixed Methods towards Health and Quality of Life* (accepted). Book chapter in: Katarzyna Wac, Sharon Wulfovich (eds.), *Quantifying Quality of Life: Incorporating Daily Life into Medicine*, Health Informatics, Springer, Cham. 30p. This publication can be found in Appendix A. [58]

    - Thoroughly explores the research area of assessing the Energy and Fatigue QoL facet in the World Health Organisation model [40] by using PROs, TechROs, and other types of reported outcomes.

4 Design and Development of the mQoL-Lab Platform

4.1 Background

During his PhD at the QoL Lab, the author of the thesis developed four modules for the equivalent of one year full-time.

4.2 Modules

The first two modules are part of the mQoL-Lab platform. The second module was also integrated into the software application of the AAL CoME European project [36], available online.

The third and fourth modules were integrated into the H2020 WellCo European project [17], available online.

1. **mQoL-Lab module for digital biomarker data collection from consumer wearables (TechRO)**. Wearable data collectors for the Quality of Life technologies lab. Implemented data collectors for most streams in the manufacturers' web APIs. Added schedulers, historic data collectors, visualisation dashboards, data exporters, and more. Collects data from Fitbit and Withings wearables. Leveraged in studies at the University Geneva, Switzerland and Stanford University, United States of America.

2. **mQoL-Lab and AAL CoME module for behavioural, health, and QoL data collection from questionnaires/validated scales (PRO)**. Implemented the data models and outcome scoring of 8 clinical instruments. Quantified risk factors: anxiety, depression, health-related life quality, memory, nutrition, physical activity, sleep, and social support. Served the questionnaires and risks as an API consumed by a web app running at HI-Iberia (the CoME coordinator partner). Released with extensive test and documentation, collected and exposed data at the University of Geneva, Switzerland.

3. **H2020 module for physical health state assessment**. Collects data from the Withings wearable API and stores it in the application's unified data store. Confronts guidelines with questionnaires and wearable data on physical activity and sleep. Identifies factors from the literature which are likely to impact the overall risk of disease. Released with extensive test and documentation, processing data continuously at HI-Iberia (the WellCo coordinator partner) in Madrid, Spain.

4. **H2020 module for chronic disease risk assessment**. Quantifies the direct risk factors for cardiovascular, pancreatic, and pulmonary disease based on medical evidence [6, 23, 59]. Generates "*if you continue like this, ...*" and "*what if?*" alternative behavioural risk scenarios. Provides the minimal behaviour changes needed to reduce the modifiable risk in each situation. Released with extensive tests (coverage 94%) and documentation, processing data at HI-Iberia.

4.3 Scientific Contributions

We leveraged the first two modules as research tools to collect the TechROs and PROs, respectively, in the study assessing the co-calibration of behavioural, health, and QoL outcomes (Chapter 3). Mads Schnoor Hansen also leveraged the first module in a minimum viable version of the mQoL-Chat chatbot that collected PRO and TechRO data as a technical contribution to his master thesis [27] assessing the motivation and facilitation of human subject study participation (Chapter 2).

While we completed the third and fourth modules in the European project on time, they did not contribute to scientific research due to participant enrollment delays beyond our control.

5 Conclusions

In this thesis, we presented our research in two areas. The first area was motivation and facilitation of human subject study participation. We reviewed the motivation factors, opportunities, and challenges to participate in human studies. We proposed a presentational model using personalised stories to improve engagement and retention of participants in human studies. We designed two unifying frameworks, one mobile app and one chatbot, to effectively conduct a wide range of human subject studies. We designed, developed, and described the mQoL-Lab platform leveraged in over ten studies. The second area was the co-calibration of behavioural, health, and Quality of Life outcomes in human subject studies. We surveyed the past work on co-calibration and reviewed the Energy and Fatigue Quality of Life facet, relevant for the human subject studies conducted in our lab. We produced the coQoL computation model to co-calibrate patient- and technology-reported outcomes. We demonstrated the robustness and feasibility of coQoL in a longitudinal, observational, in situ study assessing a cohort of 42 healthy older participants. The study reported the quality properties of the resulting data and novel patterns of relationships between physical, psychological patient-reported outcomes obtained through 8 validated scales and behavioural outcomes obtained from consumer wearables. Finally, we described the mQoL-Lab modules designed and developed by the author and their use as tools in the two research areas. The contributions in both areas can inform the design of future observational and interventional studies leveraging consumer-available technology that monitors behaviours longitudinally in situ towards assessing and improving health and Quality of Life from high-quality collected datasets.

Bibliography

[1] Mohsen Naghavi, Amanuel Alemu Abajobir, Cristiana Abbafati, et al. "Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016". In: *The Lancet* 390.10100 (2017), pp. 1151–1210 (cit. on pp. 1, 3).

[2] Reinhard Busse and Miriam Blümel. *Tackling chronic disease in Europe: strategies, interventions and challenges*. 20. WHO Regional Office Europe, 2010 (cit. on p. 1).

[3] United States Office of the Assistant Secretary for Health. "The Surgeon General's report on health promotion and disease prevention". In: *Washington, DC: US Dept of Health, Education, and Welfare* 177 (1979) (cit. on pp. 1, 9).

[4] Frank W Booth, Christian K Roberts, and Matthew J Laye. "Lack of exercise is a major cause of chronic diseases". In: *Comprehensive Physiology* 2.2 (2011), pp. 1143–1211 (cit. on p. 1).

[5] *Global recommendations on physical activity for health*. `https://www.who.int/dietphysicalactivity/global-PA-recs-2010.pdf`. Accessed: 2020-12-16 (cit. on p. 1).

[6] Peter Ueda, Mark Woodward, Yuan Lu, et al. "Laboratory-based and office-based risk scores and charts to predict 10-year risk of cardiovascular disease in 182 countries: a pooled analysis of prospective cohorts and health surveys". In: *The lancet Diabetes & endocrinology* 5.3 (2017), pp. 196–213 (cit. on pp. 1, 38).

[7] Robert Rosenman, Vidhura Tennekoon, and Laura G Hill. "Measuring bias in self-reported data". In: *International Journal of Behavioural and Healthcare Research* 2.4 (2011), pp. 320–332 (cit. on p. 1).

[8] Rodrigo Schames Kreitchmann, Francisco J Abad, Vicente Ponsoda, María D Nieto, and Daniel Morillo. "Controlling for response biases in self-report scales: forced-choice vs. psychometric modeling of Likert items". In: *Frontiers in Psychology* 10 (2019), p. 2309 (cit. on p. 1).

[9] Nancy E Mayo, Sabrina Figueiredo, Sara Ahmed, and Susan J Bartlett. "Montréal accord on patient-reported outcomes (PROs) use series–Paper 2: Terminology proposed to measure what matters in health". In: *Journal of clinical epidemiology* 89 (2017), pp. 119–124 (cit. on p. 2).

[10] ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories et al. "ATS statement: guidelines for the six-minute walk test". In: *Am J Respir Crit Care Med* 166 (2002), pp. 111–117 (cit. on p. 2).

[11] Katarzyna Wac. "Quality of Life technologies". In: *Encyclopedia of Behavioral Medicine*. Ed. by Marc Gellman. New York, NY: Springer New York, 2019, pp. 1–2 (cit. on p. 2).

[12]Andrea Coravos, Sean Khozin, and Kenneth D Mandl. “Developing and adopting safe and effective digital biomarkers to improve patient outcomes”. In: *NPJ digital medicine* 2.1 (2019), pp. 1–5 (cit. on p. 2).

[13]Jukka-Pekka Onnela and Scott L Rauch. “Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health”. In: *Neuropsychopharmacology* 41.7 (2016), pp. 1691–1696 (cit. on p. 2).

[14]Gerard Anderson. “Responding to the growing cost and prevalence of people with multiple chronic conditions”. In: *Johns Hopkins Bloomberg School of Public health* (2011) (cit. on p. 2).

[15]Joel E Cohen. “Human population: the next half century”. In: *science* 302.5648 (2003), pp. 1172–1175 (cit. on p. 3).

[16]Vlad Manea and Katarzyna Wac. “WellCo: Wellbeing and Health Virtual Coach”. In: *ERCIM News. Special Theme: Digital Health* 118 (2019), pp. 37–38 (cit. on p. 9).

[17]*Wellbeing and Health Virtual Coach - H2020 WellCo*. `http://wellco-project.eu`. Accessed: 2020-12-20 (cit. on pp. 9, 37).

[18]*International Workshop on Longitudinal Data Collection in Human Subject Studies*. `https://ldc2019ubicomp.wordpress.com`. Accessed: 2020-11-18 (cit. on p. 10).

[19]*UbiComp/ISWC ’19 Adjunct: Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. London, United Kingdom: Association for Computing Machinery, 2019 (cit. on p. 10).

[20]Vlad Manea, Allan Berrocal, Alexandre De Masi, et al. “LDC’19: international workshop on longitudinal data collection in human subject studies”. In: *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 2019, pp. 878–881 (cit. on pp. 10, 16, 23, 63).

[21]Allan Berrocal, Vlad Manea, Alexandre De Masi, and Katarzyna Wac. “mQoL Lab: Step-by-step creation of a flexible platform to conduct studies using interactive, mobile, wearable and ubiquitous devices”. In: *Procedia Computer Science* 175 (2020), pp. 221–229 (cit. on pp. 10, 24, 68).

[22]*Personas*. `https://www.usability.gov/how-to-and-tools/methods/personas.html`. Accessed: 2020-11-19 (cit. on p. 10).

[23]Jaana Lindström and Jaakko Tuomilehto. “The diabetes risk score: a practical tool to predict type 2 diabetes risk”. In: *Diabetes care* 26.3 (2003), pp. 725–731 (cit. on pp. 13, 38).

[24]Katarzyna Wac. “From quantified self to quality of life”. In: *Digital Health*. Springer, 2018, pp. 83–108 (cit. on p. 14).

[25]Riccardo Miotto, Fei Wang, Shuang Wang, Xiaoqian Jiang, and Joel T Dudley. “Deep learning for healthcare: review, opportunities and challenges”. In: *Briefings in bioinformatics* 19.6 (2018), pp. 1236–1246 (cit. on p. 16).

[26]Elisabeth Brinth Refstrup, Alba Kejser Perez, and Cecilie Rosentoft. “Designing for participation in longitudinal health and well-being studies”. MA Thesis. University of Copenhagen, 2020 (cit. on p. 19).

[27]Mads Schnoor Hansen. “Engaging participants in the recruitment phase of human subject health studies – mQoL-chat: a chatbot approach”. MSc Thesis. University of Copenhagen, 2019 (cit. on pp. 19, 38).

[28]*ResearchKit*. `https://developer.apple.com/researchkit`. Accessed: 2020-11-19 (cit. on pp. 19, 250).

[29]*HealthKit*. `https://developer.apple.com/healthkit`. Accessed: 2020-11-19 (cit. on pp. 19, 250).

[30]*Chatbot.com*. `https://www.chatbot.com`. Accessed: 2020-12-05 (cit. on p. 19).

[31]*Apple Research App*. `https://www.apple.com/ios/research-app`. Accessed: 2020-12-20 (cit. on pp. 21, 249–252).

[32]*Google Health Studies*. `https://health.google/for-everyone/health-studies`. Accessed: 2020-12-20 (cit. on p. 21).

[33]Vlad Manea and Katarzyna Wac. “mQoL: mobile quality of life lab: from behavior change to QoL”. In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 2018, pp. 642–647 (cit. on pp. 21, 23, 56, 249–252).

[34]*Google Health Studies on Google Play*. `https://play.google.com/store/apps/details?id=com.google.android.apps.health.research.studies`. Accessed: 2020-12-20 (cit. on pp. 21, 249–252).

[35]Vlad Manea, Mads Schnoor Hansen, Semahat Ece Elbeyi, and Katarzyna Wac. “Towards personalizing participation in health studies”. In: *Proceedings of the 4th International Workshop on Multimedia for Personal Health & Health Care*. 2019, pp. 32–39 (cit. on pp. 23, 47).

[36]*Caregiver and Me - AAL CoME*. `http://come-aal.eu/`. Accessed: 2020-12-20 (cit. on pp. 26, 37).

[37]Ronald Freeze and Robyn L Raschke. “An assessment of formative and reflective constructs in IS research”. In: (2007) (cit. on p. 26).

[38]HCW De Vet, CB Terwee, LB Mokkink, and DL Knol. “Concepts, theories and models, and types of measurements”. In: *Measurement in Medicine: A Practical Guide. Cambridge: Cambridge University Press* (2011), pp. 7–29 (cit. on p. 26).

[39]Gill Livingston, Andrew Sommerlad, Vasiliki Orgeta, et al. “Dementia prevention, intervention, and care”. In: *The Lancet* 390.10113 (2017), pp. 2673–2734 (cit. on p. 26).

[40]*WHOQOL: Measuring Quality of Life*. `https://www.who.int/tools/whoqol`. Accessed: 2020-10-16 (cit. on pp. 26, 36).

[41] *International Physical Activity Questionnaire (IPAQ)*. https://google.com/site/theipaq. Accessed: 2020-10-16 (cit. on p. 26).

[42] Gregory D Zimet, Nancy W Dahlem, Sara G Zimet, and Gordon K Farley. "The multidimensional scale of perceived social support". In: *Journal of personality assessment* 52.1 (1988), pp. 30–41 (cit. on p. 26).

[43] David Goldberg, Keith Bridges, Paul Duncan-Jones, and David Grayson. "Detecting anxiety and depression in general medical settings." In: *British Medical Journal* 297.6653 (1988), pp. 897–899 (cit. on p. 26).

[44] Ramón Estruch, Emilio Ros, Jordi Salas-Salvadó, et al. "Primary prevention of cardiovascular disease with a Mediterranean diet". In: *New England Journal of Medicine* 368.14 (2013), pp. 1279–1290 (cit. on p. 26).

[45] Ramón Estruch, Emilio Ros, Jordi Salas-Salvadó, et al. "Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts". In: *New England journal of medicine* 378.25 (2018), e34 (cit. on p. 26).

[46] *Mini Nutritional Assessment (Self MNA) for adults 65 years of age and older*. https://www.mna-elderly.com. Accessed: 2020-10-16 (cit. on p. 26).

[47] Alan Sunderland, Kathleen Watts, Alan D Baddeley, and John E Harris. "Subjective memory assessment and test performance in elderly adults". In: *Journal of Gerontology* 41.3 (1986), pp. 376–384 (cit. on p. 26).

[48] Janet S Carpenter and Michael A Andrykowski. "Psychometric evaluation of the Pittsburgh sleep quality index". In: *Journal of psychosomatic research* 45.1 (1998), pp. 5–13 (cit. on p. 26).

[49] *EQ-5D-3L*. https://euroqol.org/eq-5d-instruments/eq-5d-3l-about. Accessed: 2020-10-16 (cit. on p. 26).

[50] *Fitbit*. https://www.fitbit.com. Accessed: 2020-10-16 (cit. on p. 26).

[51] Mary E Rosenberger, Janet E Fulton, Matthew P Buman, et al. "The 24-hour activity cycle: a new paradigm for physical activity". In: *Medicine and science in sports and exercise* 51.3 (2019), p. 454 (cit. on p. 27).

[52] Kirke Kjellberg. "ConsistencyQoL – A framework for modelling consistency in behavioural data collected with wearables". MSc Thesis. University of Copenhagen, 2020 (cit. on p. 30).

[53] Sharon Wulfovich, Maddalena Fiordelli, Homero Rivas, Waldo Concepcion, and Katarzyna Wac. ""I must try harder": Design implications for mobile apps and wearables contributing to self-efficacy of patients with chronic conditions". In: *Frontiers in psychology* 10 (2019), p. 2388 (cit. on p. 31).

[54] *Vital Beats*. https://vitalbeats.com. Accessed: 2020-12-18 (cit. on p. 33).

[55] Vlad Manea and Katarzyna Wac. "Co-Calibrating physical and psychological outcomes and consumer wearable activity outcomes in older adults: An evaluation of the coQoL method". In: *Journal of Personalized Medicine* 10.4 (2020), p. 203 (cit. on pp. 35, 93).

[56]Vlad Manea, Allan Barrocal, and Katarzyna Wac. “Using consumer-friendly wearables to correlate patient and technology-reported physical activity in healthy seniors”. In: *Procedia Computer Science* 175 (2020), pp. 245–252 (cit. on pp. 35, 180).

[57]“27th Annual Conference of the International Society for Quality of Life Research”. In: *Quality of Life Research* 29.1 (Oct. 2020), pp. 1–196 (cit. on pp. 36, 189).

[58]Natalie Solomon and Vlad Manea. “Quantifying energy and fatigue: Classification and assessment of energy and fatigue using subjective, objective, and mixed methods towards health and Quality of Life”. In: *Quantifying Quality of Life: Incorporating Daily Life into Medicine*. Ed. by Katarzyna Wac and Sharon Wulfovich. Cham, Switzerland: Springer, Cham, 2021 (cit. on pp. 36, 201).

[59]Fernando J Martinez, Anastasia E Raczek, Frederic D Seifer, et al. “Development and initial validation of a self-scored COPD Population Screener Questionnaire (COPD-PS)”. In: *COPD: Journal of Chronic Obstructive Pulmonary Disease* 5.2 (2008), pp. 85–95 (cit. on p. 38).

Publications

A

This appendix contains the publications included in this PhD thesis. A full list of publications is available in my academic curriculum vitae (Appendix B.4).

Motivation and Facilitation of Human Subject Study Participation

Publication 1: Towards Personalizing Participation in Health Studies

Vlad Manea, Mads Schnoor Hansen, Semahat Ece Elbeyi, Katarzyna Wac. *Towards Personalizing Participation in Health Studies*. Workshop on Multimedia for Personal Health and Health Care (HealthMedia 2019), Conference on Multimedia (MM 2019). 8p. DOI: `https://doi.org/10.1145/3347444.3356241`. [35]

Abstract There is substantial evidence on the relevant factors that motivate participation in human subject studies and the expectations of participants when sharing their health data for research. However, most human subject studies focus on participant eligibility and data collection, omitting even a rudimentary use of the factors that motivate participation. We illustrate an approach to use motivation to construct personalized stories and exemplify it by using a chatbot under development towards monitoring, analyzing, and influencing health study participation, engagement, and retention. Additionally, we discuss the new advantages, challenges, and unexplored avenues for research stemming from our approach.

Keywords health study, participation motivation assessment, chatbot.

Towards Personalizing Participation in Health Studies

Vlad Manea
University of Copenhagen
manea@di.ku.dk

Mads Schnoor Hansen
University of Copenhagen
qdp777@alumni.ku.dk

Semahat Ece Elbeyi
University of Copenhagen
ece.elbeyi@hum.ku.dk

Katarzyna Wac
University of Geneva
University of Copenhagen
katarzyna.wac@unige.ch
wac@di.ku.dk

ABSTRACT

There is substantial evidence on the relevant factors that motivate participation in human subject studies and the expectations of participants when sharing their health data for research. However, most human subject studies focus on participant eligibility and data collection, omitting even a rudimentary use of the factors that motivate participation. We illustrate an approach to use motivation to construct personalized stories and exemplify it by using a chatbot under development towards monitoring, analyzing, and influencing health study participation, engagement, and retention. Additionally, we discuss the new advantages, challenges, and unexplored avenues for research stemming from our approach.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**; *Empirical studies in ubiquitous and mobile computing*.

KEYWORDS

Health Study, Participation Motivation Assessment, Chatbot

ACM Reference Format:
Vlad Manea, Mads Schnoor Hansen, Semahat Ece Elbeyi, and Katarzyna Wac. 2019. Towards Personalizing Participation in Health Studies. In *4th International Workshop on Multimedia for Personal Health and Health Care (HealthMedia '19), October 21, 2019, Nice, France.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3347444.3356241

1 INTRODUCTION

Significant work foregrounded several areas (factors) of motivation to participate in health studies: *seriousness* (trustworthiness of the study and the researchers, data security, data protection agency approvals, ethics review board assessment), *altruism* (to support family, friends, community, and society), *personal benefits* (study results, intellectual curiosity, financial gains), *study-related reasons* (effort to participate, means of participation, time invested, data required), and *external motivation* (participation of others, media reports, endorsements from official institutions) [2].

However, health studies notoriously suffer from limited enrollment, engagement, and retention. Across the fragmented universe of specialized health studies, study researchers analyze participation through retroactive analyses of the participant (e.g., demographics, availability, eligibility, health state), the study (e.g., topic, duration), and the participation (e.g., attrition segmented by demographics and topic), with no assessment of the motivation to participate [9]. The limited understanding of the motivation to participate throughout the study leads to decreased *engagement* (e.g., performing tasks, answering questions, results reflection) and *retention* (e.g., wearing a personal wrist device specifically for the study for a given duration), with adverse effects on the quality of the collected data.

We argue that motivation-driven participation can be monitored, analyzed, and (positively) influenced from the moment of enrollment, throughout the study, and up to the abandonment or completion (whichever occurs first). The study can provide the participant with a personalized *story*, stemming from the assessment of motivation. A story can be represented as a sequence of *moments*, characterized by relevant informational content embedded in the regular user interface. Moments could reside in not only the existing medium of the study (e.g., a text field in a form, a video to vote, a push notification to swipe in a mobile app, an interactive task to execute on a web user interface, a chart for wearable data visualization, a conversation message in a chatbot). They could also use the context of personal motivation. For example, a question could be preceded by a message relevant exclusively for the participant and the question. Another example is the visualization of study results followed by an explanation of how these results apply directly to the participant. Other examples include the tangible impact of their contributions on society upon attrition-preventing notification, the announcement of financial gains before/after participation, and new features to spark personal curiosities when engagement is low.

Assessing the degree to which a story personalizes participation and creates a sense of belonging for the participant, more than the straightforward instructions about the actions to perform, is an unexplored area of research, with the potential to improve enrollment, engagement, and retention rates, and increase the quality of the resulting collected data.

This article is structured as follows. Section 1 provided an introduction to assessing motivation for participation. Section 2 reviews the related work in this research area. Section 3 familiarizes the

HealthMedia '19, October 21, 2019, Nice, France

ACM ISBN 978-1-4503-6914-5/19/10…$15.00
https://doi.org/10.1145/3347444.3356241

reader with a case study. Section 4 introduces our approach and applies it to the case study. Section 5 describes our plan to assess participation motivation by operationalizing a chatbot and constructing participation models from the collected data. Section 6 highlights the future work, and Section 7 concludes the article.

2 RELATED WORK

Previous work assessed the self-reported [22] willingness to participate in health studies from the perspective of sharing electronic health record and wearable data with researchers as well as the reasons for people to participate in health studies.

2.1 Willingness to Participate

In general, participants are more willing to participate in health studies and share their data with researchers than other institutions outside of the circle of health care providers. While participants favor more sharing their complete data with their primary care provider, they agree to share specific parts of their data with researchers given the provided data has a particular use, and they keep ownership of the data.

Sharing Electronic Health Records Perera et al. (2011) asked 490 diabetes patients in Toronto, ON, Canada to assess perceived benefits and harms of keeping an Electronic Health Record (EHR) on a 5-level Likert scale. 67% supported the use of their de-identified data by researchers [26]. Caine et al. (2012) asked 30 health care recipients in Bloomington, IN, the United States to pair parts of the EHR with potential recipients, including researchers. Only 10% of those with sensitive EHR parts were willing to share all EHR parts with researchers, and only 15% were willing to share all non-sensitive EHR parts with them. However, the participants were willing to anonymously and temporarily share relevant parts of their EHRs with researchers [4].

Sharing Wearable Health Data Bietz et al. (2015) surveyed 465 individuals from San Diego, CA, the United States on the challenges of sharing wearable health data with research. 78% of the participants were willing to anonymously donate their data for research, especially if it would contribute to the greater good, and would favor exciting over noninteresting studies [1]. Seifert et al. (2018) interviewed 1.013 older age participants from Zurich, Switzerland, and found that 57% would be willing to share health data with researchers [28]. Chen et al. (2016) studied the mobile and wearable fitness tracking behaviors of 101 people from Sydney, Australia, and reported that 77% were willing to donate health data [5].

2.2 Motivation to Participate

In general, altruistic motivation and personal benefits overshadow health gains and financial incentives, and this trend pronounces recently. Participants are at first motivated by helping others (from family to society) and then, over time, by the personal benefits from the studies (relevant results, intellectual curiosity, improving health). Trust (research institute, data protection, ethics review) is fundamental for the expression of these motivation factors.

Healthy Populations Trauth et al. (2010) surveyed 489 healthy people in Southwestern Pennsylvania, PA, United States and found that 46% of the respondents would take part in a health study focusing on curing a specific disease of interest. In particular, having a sick friend or relative contributed to the willingness to participate [32]. Stunkel and Grady (2011) reviewed the motivations of 2.000 healthy people participating in 12 English-speaking clinical trials between 1977 and 2010 in the United States, Europe, and Africa. They found that financial motivation was one of the main motivations in the studies. However, helping to contribute to science and medicine, helping others, and taking part in "something important" were also reported [30]. Kerath et al. (2013) reviewed 1.041 healthy people's beliefs and attitudes towards participating in genetic research through a 22-item questionnaire distributed in a network of hospitals in Long Island, NY, United States. 83% considered participation important for society. 82% would approve genetic research, 70% would be willing to participate anonymously in research studies with genetic data, and 53% would be willing to participate in a named biobank study. Among those not willing, 74% would refuse due to data privacy concerns. Other general concerns include enrolling in a study not well explained (over 61%), in additional research conducted without their knowledge (over 74%), or in other kinds of research without their knowledge (over 62%) [18]. Nobile et al. (2017) found from 623 questionnaire respondents from two German studies that a contribution to society instead of personal benefits appears to contribute to participant retention [24]. Bongartz et al. (2017) asked 135 healthy people from Hannover, Germany, to self-report on a 5-level Likert scale the importance of 16 reasons grouped in four areas. They were, in the order of importance, the seriousness of the study and altruism (most important), study-related, personal benefits, and external motivation (least significant). Important reasons were the seriousness (of the study, of the researchers, of the institution), the possibility to support the researchers, and a data protection approval. Feedback of results and an exciting topic were of lesser importance. Reports in the media, the participation of friends, and financial incentives were among the least essential [2]. In a review comprising motivation factors for genetic study participation from over 6.000 healthy people, Goodman et al. (2018) found the financial compensation the least important [12]. Kim et al. (2019) surveyed 170 senior South Koreans and found that 39% would share health information with researchers, below family and hospitals, and above device manufacturers, insurance companies, and governmental agencies [19].

Diseased Populations McCann et al. (2010) conducted a qualitative study with 13 heartburn patients and reported that participants contribute to trials for the greater good initially, but sustained participation over time depends on a concrete personal benefit and no significant personal disadvantages (weighing risks and benefits), a term called *conditional altruism* [23]. Soule et al. (2016) asked 164 cardiac patients in Boston, MA, the United States to self-report their agreement with four areas of reasons after participating in a study on a 10-level Likert scale. The areas were, in the order of reported importance: altruistic (most important), intellectual, health-related, and financial (least relevant) [29]. Goodman et al. (2019) asked 450 cancer patients, controls, and relatives in the West Washington area, WA, the United States to self-report the importance of several areas of reasons for participating in a 4-level Likert scale. Resulting important themes include benefits for society (highly significant overall), the reputation of institution (highly important overall and particularly mentioned in the context of data collection), benefits

for family and known people (both important), research meaningfulness personally (particularly important for cancer cases), and a much less crucial financial incentive [11].

We did not observe work on quantifying the *say versus do* relationship between self-reported participant motivation (above), self-reported personality traits (assessed with validated scales, such as the Big Five test for personality traits [10]), objective study attributes, and technology-reported [22] participation on one end, and the outcomes of participation (enrollment, engagement, retention) on the other end.

3 CASE STUDY

An active area of research is the quantification of the modifiable risk of chronic disease at an older age by observing (modifiable) daily life behaviors at a younger age in healthy populations. For example, physical activity is a direct risk factor for multiple types of neoplastic [6] and metabolic [20] diseases, and an indirect risk factor for a wide range of chronic diseases [3].

3.1 Study Setup

In our research lab, an observational study (on chronic disease risk assessment based on the physical activity daily life behaviors) quantifies the relationship between the continuous physical activity (types, intensities, calories, distances, durations, and more) technology-reported by a wearable device (e.g., Fitbit [13]) during daily living and the physical activity self-reported *a posteriori* by the participant through a validated questionnaire. One such questionnaire is the International Physical Activity Questionnaire (IPAQ, [7]) which assesses levels of physical activity intensity (sedentary, walking, moderate, vigorous) for the past seven days. The study assesses physical activity by discovering behavioral patterns of daily life unfolding in time and context, without burdening the participant to simultaneously wear for comparison a second clinically validated device (e.g., ActiWatch [25]).

A strong relationship between the wearable- and questionnaire-reported intensities of physical activity would suggest that this risk factor could be estimated by using the consumer wearable. Fluctuations in the physical activity level over short-term periods would yield to changes in the risk over long-term periods.

3.2 Study Participants

The participants are three people who own wearable devices and are willing to enroll in health studies.

Alice is a mother with genetic diabetes risk. She wants to help her mother Edna cope with diabetes, herself to monitor her metabolism, and her daughter to avoid this disease altogether. Alice puts her trust in research for the greater good, and her participation does not depend on financial incentives.

Bob is a hardware programmer during the day, and a Quantified Self [31] tracker during the night. He enrolls in studies which challenge the understanding of his behaviors and experiments with the latest wearable features.

Charlie is a student on a mission to find more money to cover rent and tuition in a big city. He is invited to many research studies on campus and wants to use this in his favor. He enrolls in as many studies as possible but invests as little time as possible in each.

4 PARTICIPATION STORIES

From a participant experience perspective, at enrollment time, the study would introduce the scientific purpose and research institution, enumerate the actions required by the participant, request the legal consent, and ask for permission to unobtrusively access the physical activity data from the participant's wearable device.

4.1 Uniform Participation Story

Throughout the study (months to years), in its strictest setup, the elapse of any contiguous seven day period with continuously collected physical activity wearable data would trigger an ecological momentary assessment [33] of the participant to assess own physical activity for the same period by responding to the questionnaire. More likely, the study would relax this requirement and administer the questionnaire more rarely, at moments seven days apart or upon schedule (e.g., every Sunday evening).

In a basic setup, the interactions with each participant would be concise, focused on the collected data. However, they would also be uniform across all participants, as they would lack personalization. An example in the context of a chat conversation-based study appears by Figure 1. After several weeks of contribution, the participants will forget the goal of the research and the rationale for joining it in the first place. In the absence of any personalized interaction with the study, they will lose interest and abandon the investigation.

4.2 Personalized Participation Stories

In an alternative setup, each participant is asked about the motivating factors to participate in health studies in general. For example, a set of questions on the factors of motivation may be administered at the moment of enrollment in the first study, soon afterward, or split across the registrations to an early few studies deployed in a shared medium. In the example, the questions are administered before signing up for the first study.

Alice specifies she is motivated to help others, to gain personal benefits, and by aspects in the study method. Within helping others, she wishes to help her parent, her child, people like her, and future generations. From the possible personal benefits, Alice is only interested in finding out if she respects the guidelines. In a study, she is comfortable with answering questions from home in a chatbot but is worried about the security of the data she shares. Figure 2 depicts her choices in a possible set of questions. At enrollment time, these factors determine Alice surpass the risk-benefit analysis in favor of the benefits of participation. For Alice, her mother, her daughter, and she can act as motifs of the story, on the background of reinforced total control and ownership of the data. Although the study follows a standard data management protocol for all participants, the study clarifies for Alice the protocol (e.g., a simple sequence of steps), the data flow (e.g., a diagram with the entities and channels of data transfer from wearable to research study storage), the enforced data security properties (e.g., anonymity and privacy, with simple explanations), privacy statement (summary) and terms of use (summary), ethics review board approvals, etc. A moment in Alice's participation story, containing a selection of the information above, is depicted in Figure 3.

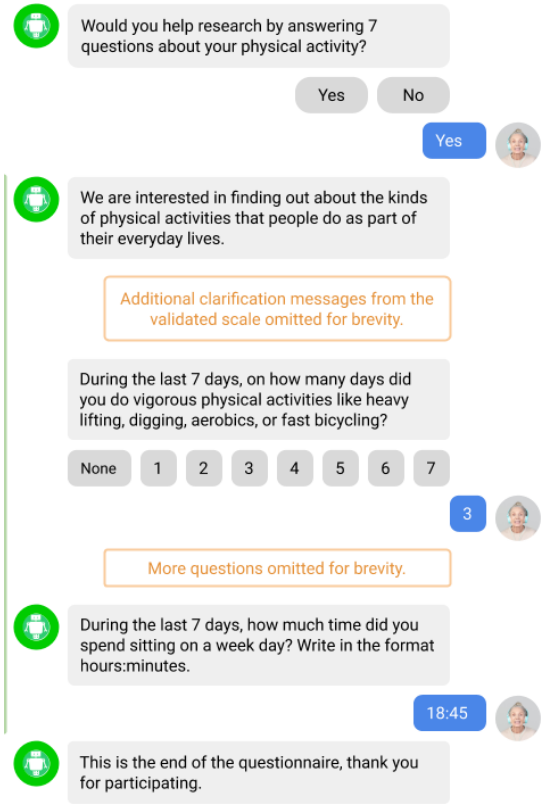


Figure 1: Uniform Story. The study interacts with each participant within no story. While concise, this interaction does not personalize participation. The pale green line indicates the administration of the validated scale, characterized by little flexibility in the presented information.

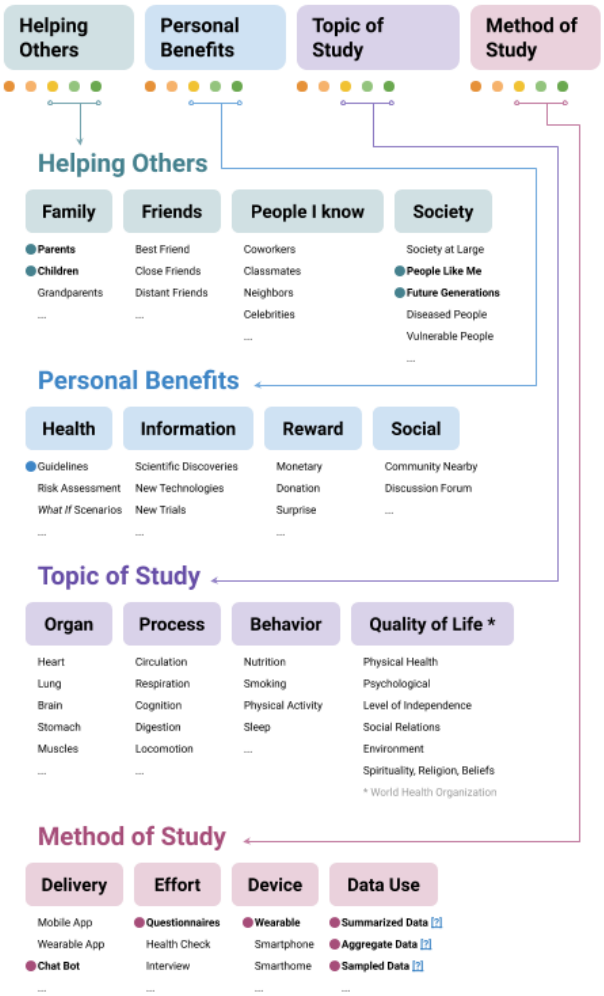


Figure 2: Motivation Factors Selection by Alice. The selection assesses each participant's motivation factors. First, the participant rates the level of motivation for each of the four areas. Then, for the areas which highly motivate, moderately motivate, and mildly motivate participation, the person can select more specific factors of motivations. The selection resembles the areas and factors from our literature review.

Bob is primarily interested in personal benefits: scientific discoveries and new technologies, as well as health guidelines. Then, he is interested in the study method, specifically in using his wearable. Similar to Alice's case, he is notified to provide a momentary assessment of personal physical activity, as depicted in Figure 4, and the validated scale parts of the conversation are identical. However, the presentation of his story focuses on wearable technology rather than data security. It shows the data collected and the outcome of the measurement in technical terms. However, it does not detail the security properties of the data collected.

Charlie is laser-focused on the reward type of personal benefits and prefers internet use, but is wondering whether other students like him participate. In Figure 5, his story focuses on the financial reward and the opportunity of meeting like-minded people.

5 ASSESSING MOTIVATION

Our lab models current behaviors towards preventing the onset of disease in the future. While experienced in mobile development and data analysis from longitudinal daily life wearable data, we have difficulties in recruiting participants for our studies, such as the new study on physical activity. For example, in the case of the physical activity study, we recruited 18 participants and collected only 35 observations leading to moderate to high correlations, however only on a minimal palette of physical activity intensities. The scarcity of participants drove us to take one step back and investigate the factors, challenges, and opportunities that affect the motivation to participate in health studies in the first place.

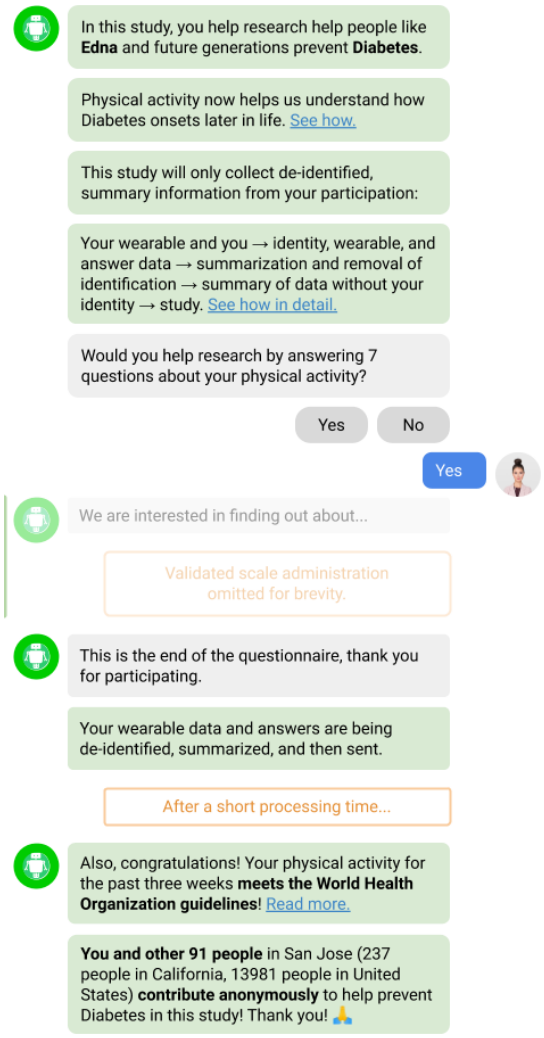


Figure 3: Personalized Story for Alice. The study interacts with Alice through a story. While less concise, this interaction personalizes participation with moments (green).

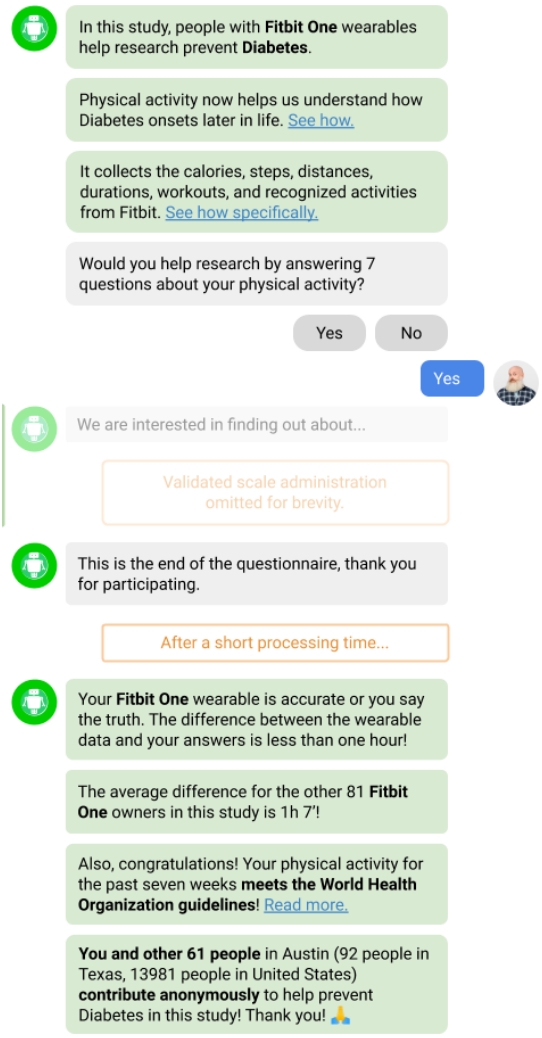


Figure 4: Personalized Story for Bob. The study interacts with Bob through a story. While less concise, this interaction personalizes participation with moments (green).

5.1 Building a Chatbot for Studies

To reach a wider audience, first to assess motivation and then to enroll them into our studies, we are operationalizing a chatbot. We found the chatbot to be the most appropriate (and yet lesser used in the past) medium for conducting our research. The chatbot is designed to assess motivation before the enrollment in a new study, enroll participants in the first study for physical activity, and monitor the participant retention and engagement for an initial understanding of the motivation of participation. The chatbot can be deployed onto multiple *platforms* (currently, Messenger [14] and our website), and its conversation capabilities are implemented with a third party *builder* (ChatBot.com [16]). An example of interaction with the chatbot is depicted in Figure 6.

5.2 Advantages

We have identified numerous practical advantages for participants, studies, and ourselves as developers by using a chatbot.

5.2.1 Easy Access for Participants. The Chatbots can be initiated by navigating to an internet address and opening a conversation in the favorite messaging system of the participant, available and keeping its state across devices (e.g., phone, tablet, and desktop), without downloading a mobile application.

5.2.2 Support for Participants. Chatbots allow for human intervention in the chat conversations. This fact allows us as researchers to resolve usability issues faster for early adopters.

5.2.3 Natural Language for Participants. From text fields filled with natural language, the chatbot builder can automatically extract

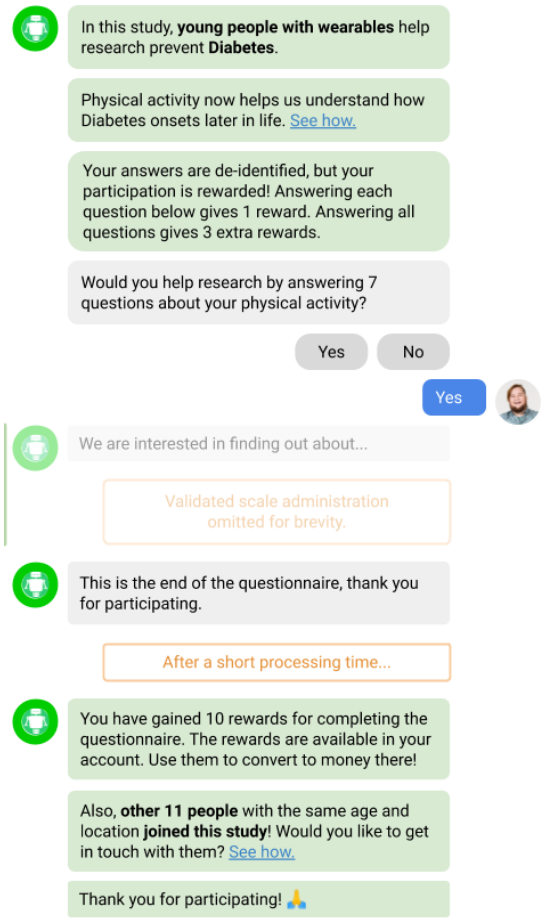


Figure 5: Personalized Story for Charlie. The study interacts with Charlie through a story. This interaction personalizes participation with moments (green).

specific concepts (e.g., the topics of studies in which a participant is interested) for further processing.

5.2.4 Flexibility for Studies. The user interface controls available in the conversation (e.g., images, videos, text, links, text fields, buttons) are sufficiently expressive for most studies. For more advanced features (e.g., video recording, sound recording, location monitoring), the chatbot can show in the conversation a responsive web window to a website that implements them. The chatbot can render a web window to a website for wearable signup in the same way.

5.2.5 Social Networks Leverage for Studies. The chatbot cannot participate in conversations with more people. However, it can unite two people through their discussions with it.

5.2.6 Programmatic Specification for Developers. For Developers, the chatbot can be specified as code, reducing the risk of errors when constructing it (traditionally, by navigating visual flows of parts of conversations in a web dashboard) and opening the possibility for automated testing and version control.

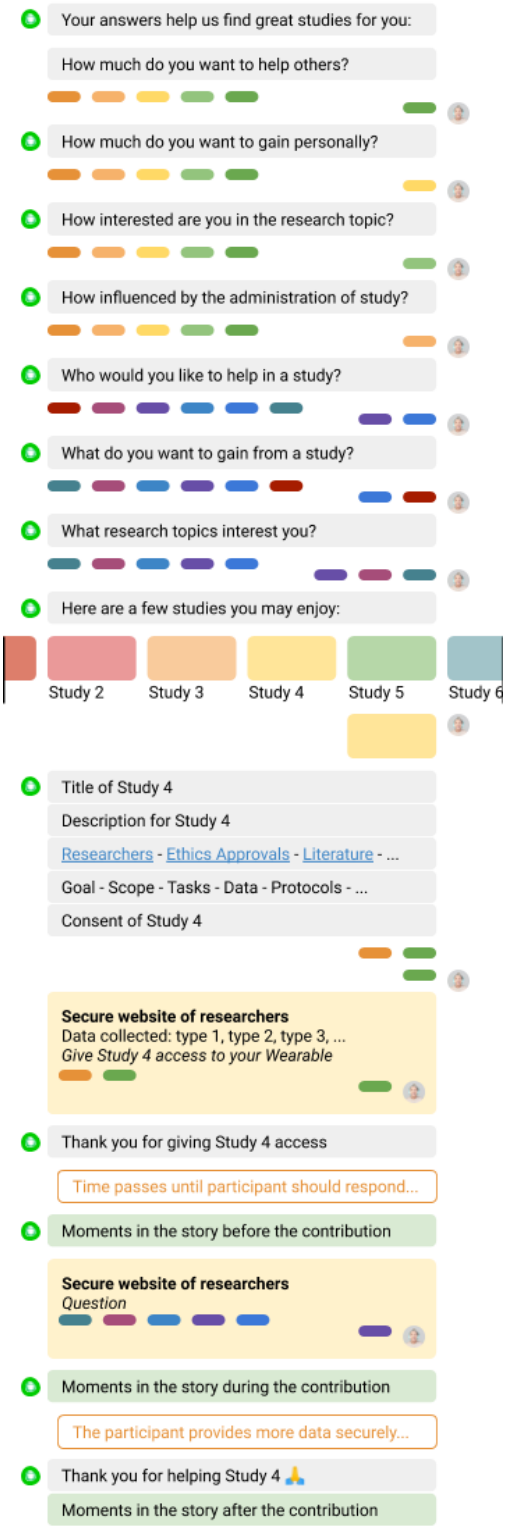


Figure 6: Interaction with the Chatbot. Assessment of motivation areas and factors, selection and enrollment in a study, secure wearable data signup, and contribution to the study (yellow), story moments (green).

5.2.7 Rapid Prototyping for Developers. The chatbot is adequate for prototyping and experimentation in the nascent phase of our studies. It allows developers to add, modify, and remove parts of the conversation while participants use it.

5.2.8 Bidirectional Communication for Developers. The chatbot also allows bidirectional communication with a server through a secure transfer protocol, allowing developers to integrate with usage monitoring platforms (e.g., Mixpanel [17]) and create the next user interface controls in a conversation.

5.2.9 Noise Reduction for Motivation Assessment. When comparing different moments for different groups of participants, we consider a simple chatbot user interface control consistent.

5.3 Challenges

Along with the advantages, we have also identified several challenges, for which we wish to open for discussion possible approaches to reduce their impact.

5.3.1 Content Creation. Motivation-specific content can be vast: success stories, general statistics showing global scale, statistics showing local engagement, getting in touch with someone who succeeded in one's city, meet the researchers (in one's city), enroll one's granny too, participate in similar studies, and others. They can lead to a curated list of content proposed for researchers.

Upon creating a study, researchers may need to provide content related to their studies continuously: endpoint to showcase results (e.g., new paper published and its impact), new advances in the area of the study (e.g., new gene found), further clinical trials, financial incentives (based on participation, anticipation, variable surprise). While all such content requires involvement from study owners, we argue that the deployment of such updates is more straightforward than deploying them from scratch.

5.3.2 Selection Bias. The additional dimension of motivation to guiding participants towards and through studies may lead to further fragmentation of the population of interest, potentially adding to both self-selection bias (those people who are willing to use a chatbot and consider the chatbot a trustworthy enough media for a research study) and technology-selection bias (people see only studies which match their interests as assessed by our understanding about them).

An approach to reduce the latter is to balance the studies which relate to the expressed areas of interest and factors of motivation with studies which relate less, but have a high potential impact (e.g., researching a widespread disease such as one of the top worldwide killers [27], a condition relevant for the community of the participant, or a situation for which the participant's demographics are factors of risk). For short-term studies, population class imbalance due to missing representatives sharing similar demographics with the participant is an opportunity for impactful participation.

5.3.3 Self-Reported Motivation Data Provision. Figure 2 depicts one of the possible orderings for surveying the motivation factors and only one grouping of the areas and factors. However, some participants may have noncommunicable preferences for other arrangements. For example, some participants may prefer to answer all questions at the start, others may prefer to respond immediately after enrolling in a study, others may prefer a separation (e.g., to answer the topics of interest, then enrol for a study, and then clarify the other motivation factors), others may be willing to specify their preferences through ecological momentary assessments, and more.

This challenge adds a dimension to the feasibility of collecting the motivation factors of the participants while minimizing any adverse effect onto the study participation (e.g., too many questions about motivation influence the participant to drop the physical activity study faster).

5.3.4 Data Security. Currently, many chatbot platforms and builders only encrypt the transfer channels: from the participant to the chatbot platform, and from the platform to the builder.

One temporary approach to protecting data is to collect sensitive data through the web windows in the conversation (e.g., answers about health data, enrollment for wearable data collection), at the price of input method fragmentation for participants.

The builder plans to implement integrations with additional platforms (e.g., WhatsApp [15]) that allow the end to end encryption of all messages between the participant and the builder. The use of such a platform would disallow it from accessing the contents.

5.4 Learning from Participation

Long-term participation creates the opportunity to construct participation models, using as input baseline properties of the participant (e.g., demographics, topics of interest, factors, and concerns for participation, reasons for abandonment), initial features of study (e.g., goal, story, requirements, rewards), and properties of involvement (e.g., responses to moments as markers for engagement, data collection as marker for retention, outcome of completion or attrition) that change in time.

A supervised learning model for participation can predict as output the likelihood of a person with similar properties to stay in a study for a given duration. Another model can use favorable/adverse reactions to moments as rewards/punishments for reinforcement learning towards selecting the story that maximizes retention/engagement for a given period forward. The creation of models such as these is, to our knowledge, an unexplored area of research to assess study participation motivation.

6 FUTURE WORK

Along 2019 we expect to recruit 60 participants in the physical activity study from the H2020 WellCO project [21] with study participants of senior age from Denmark, Spain, and Italy (30 participants) and the Sport & Wellbeing & Health Survey [8] with young and athletic participants from Denmark (30 participants). They will be invited to participate in the motivation assessment upon enrollment in the physical activity study. Some of them are located in the same city as ourselves, allowing us to assess the feasibility of both studies more easily (e.g., meet with them in person).

We are building the chatbot (one of the authors is writing his master thesis on this topic). In its first iteration, the chatbot will implement the physical activity study and operationalize the motivation assessment for the participants who will have accepted our invitation above. The bot will initially collect self-reported demographic properties, motivation factors, and personality traits, towards creating a baseline supervised learning model to predict

the time to a participation event (engagement, retention), first without administering interventions (personalized moments) in the case study. Further, depending on the baseline results, we will experiment with responses towards more advanced modeling.

We trust the HealthMedia workshop can help us improve our preliminary research. We want to exchange experiences from other researchers conducting human subject health studies. We are interested in further our understanding of the human and technological factors of motivation to participate, as well as operational considerations learned from practical experience, that can ultimately impact participation (and how can they be measured feasibly, unobtrusively, and reliably). Then, we are interested in learning general experiences from other studies operationalized less with purpose-built mobile applications, and more on media where the participant already spends a significant amount of time, such as the messenger.

7 CONCLUSION

Although substantial research foregrounded self-reported factors that motivate participation in health-related human subject studies, little research has been done to quantify the relationship between motivation factors, personality traits, study properties, participation experience, and participation outcomes (enrollment, engagement, and retention). We proposed an approach that constructs personalized stories of participation with moments embedded in the interactions of the study. Motivated by its opportunities, we have chosen a chatbot to implement our studies and perform a preliminary quantification of the relationship above, not without challenges.

ACKNOWLEDGMENTS

EU H2020 WellCO project (769765): http://wellco-project.eu

REFERENCES

[1] Matthew J Bietz, Cinnamon S Bloss, Scout Calvert, Job G Godino, Judith Gregory, Michael P Claffey, Jerry Sheehan, and Kevin Patrick. 2015. Opportunities and challenges in the use of personal health data for health research. *Journal of the American Medical Informatics Association* 23, e1 (2015), e42–e48.

[2] Hannah Bongartz, Nicole Rübsamen, Heike Raupach-Rosin, Manas K Akmatov, and Rafael T Mikolajczyk. 2017. Why do people participate in health-related studies? *International Journal of Public Health* 62, 9 (2017), 1059–1062.

[3] Frank W Booth, Christian K Roberts, and Matthew J Laye. 2011. Lack of exercise is a major cause of chronic diseases. *Comprehensive Physiology* 2, 2 (2011), 1143–1211.

[4] Kelly Caine and Rima Hanania. 2012. Patients want granular privacy control over health information in electronic medical records. *Journal of the American Medical Informatics Association* 20, 1 (2012), 7–15.

[5] Juliana Chen, Adrian Bauman, and Margaret Allman-Farinelli. 2016. A study to determine the most popular lifestyle smartphone applications and willingness of the public to share their personal data for health research. *Telemedicine and e-Health* 22, 8 (2016), 655–665.

[6] Graham A Colditz et al. 2000. Harvard report on cancer prevention volume 4: Harvard Cancer Risk Index. *Cancer Causes & Control* 11, 6 (2000), 477–488.

[7] Cora L Craig, Alison L Marshall, Michael Sjöström, Adrian E Bauman, Michael L Booth, Barbara E Ainsworth, Michael Pratt, ULF Ekelund, Agneta Yngve, James F Sallis, et al. 2003. International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise* 35, 8 (2003), 1381–1395.

[8] Stefano De Dominicis and Katarzyna Wac. 2019. University of Copenhagen: Sport & Wellbeing & Health Survey. https://www.surveymonkey.com/r/kusportandwellbeing. [Online; accessed 1-July-2019].

[9] Caroline Free, Gemma Phillips, Leandro Galli, Louise Watson, Lambert Felix, Phil Edwards, Vikram Patel, and Andy Haines. 2013. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS Medicine* 10, 1 (2013), e1001362.

[10] Lewis R Goldberg. 1990. An alternative "description of personality": the big-five factor structure. *Journal of Personality and Social Psychology* 59, 6 (1990), 1216.

[11] Deborah Goodman, Deborah Bowen, Catherine O Johnson, Lari Wenzel, and Karen Edwards. 2019. Factors that motivate participation in observational genetic cancer research studies. *Open Journal of Epidemiology* 9, 2 (2019), 156–172.

[12] Deborah Goodman, Deborah Bowen, Lari Wenzel, Paris Tehrani, Francis Fernando, Araksi Khacheryan, Farihah Chowdhury, Catherine O Johnson, and Karen Edwards. 2018. The research participant perspective related to the conduct of genomic cohort studies: A systematic review of the quantitative literature. *Translational Behavioral Medicine* 8, 1 (2018), 119–129.

[13] Fitbit Inc. 2019. Fitbit. Fit for everybody. https://www.fitbit.com. [Online; accessed 1-July-2019].

[14] Facebook Inc. 2019. Messenger. Be together, whenever. https://www.messenger.com. [Online; accessed 1-July-2019].

[15] Facebook Inc. 2019. WhatsApp. Simple. Secure. Reliable messaging. https://www.whatsapp.com. [Online; accessed 1-July-2019].

[16] LiveChat Inc. 2019. ChatBot. Automate customer service with ChatBot within 10 minutes. https://www.chatbot.com. [Online; accessed 1-July-2019].

[17] Mixpanel Inc. 2019. Mixpanel. Grow your business by learning why users convert, engage, and retain. https://www.mixpanel.com. [Online; accessed 1-July-2019].

[18] Samantha M Kerath, Gila Klein, Marlena Kern, Iuliana Shapira, Jennifer Witthuhn, Nicole Norohna, Myriam Kline, Farisha Baksh, Peter Gregersen, and Emanuela Taioli. 2013. Beliefs and attitudes towards participating in genetic research–a population based cross-sectional study. *BMC Public Health* 13, 1 (2013), 114.

[19] Tae Kyung Kim and Moon Choi. 2019. Older adultsâĂŹ willingness to share their personal and health information when adopting healthcare technology and services. *International Journal of Medical Informatics* 126 (2019), 86–94.

[20] Jaana Lindström and Jaakko Tuomilehto. 2003. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care* 26, 3 (2003), 725–731.

[21] Vlad Manea and Katarzyna Wac. 2019. WellCo: Wellbeing and Health Virtual Coach. *ERCIM News. Special Theme: Digital Health* 118 (2019), 37–38.

[22] Nancy E Mayo, Sabrina Figueiredo, Sara Ahmed, and Susan J Bartlett. 2017. Montreal accord on patient-reported outcomes (PROs). Paper 2: terminology proposed to measure what matters in health. *Journal of Clinical Epidemiology* 89 (2017), 119–124.

[23] Sharon K McCann, Marion K Campbell, and Vikki A Entwistle. 2010. Reasons for participating in randomised controlled trials: conditional altruism and considerations for self. *Trials* 11, 1 (2010), 31.

[24] Hélène Nobile, Pascal Borry, Tobias Pischon, Astrid Steinbrecher, Heiner Boeing, Matthäus Vigl, and Manuela M Bergmann. 2017. ParticipantsâĂŹ decision to enroll in cohort studies with biobanks: quantitative insights from two German studies. *Personalized Medicine* 14, 6 (2017), 477–485.

[25] Koninklijke Philips N.V. 2019. Philips Actigraphy ActiWatch. http://www.actigraphy.com/solutions/actigraphy.html. [Online; accessed 1-July-2019].

[26] Gihan Perera, Anne Holbrook, Lehana Thabane, Gary Foster, and Donald J Willison. 2011. Views on health information sharing and privacy from primary care practices using electronic medical records. *International Journal of Medical Informatics* 80, 2 (2011), 94–101.

[27] Gregory A Roth, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi, Hedayat Abbastabar, Foad Abd-Allah, Jemal Abdela, Ahmed Abdelalim, et al. 2018. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 392, 10159 (2018), 1736–1788.

[28] Alexander Seifert, Markus Christen, and Mike Martin. 2018. Willingness of older adults to share mobile health data with researchers. *Journal of Gerontopsychology and Geriatric Psychiatry* 31, 1 (2018), 41–49.

[29] Michael C Soule, Eleanor E Beale, Laura Suarez, Scott R Beach, Carol A Mastromauro, Christopher M Celano, Shannon V Moore, and Jeff C Huffman. 2016. Understanding motivations to participate in an observational research study: Why do patients enroll? *Social Work in Health Care* 55, 3 (2016), 231–246.

[30] Leanne Stunkel and Christine Grady. 2011. More than the money: a review of the literature examining healthy volunteer motivations. *Contemporary Clinical Trials* 32, 3 (2011), 342–352.

[31] Melanie Swan. 2013. The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data* 1, 2 (2013), 85–99.

[32] Jeanette M Trauth, Donald Musa, Laura Siminoff, Ilene Katz Jewell, and Edmund Ricci. 2000. Public attitudes regarding willingness to participate in medical research studies. *Journal of Health & Social Policy* 12, 2 (2000), 23–43.

[33] Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2018. The experience sampling method on mobile devices. *Comput. Surveys* 50, 6 (2018), 93.

Publication 2: Mobile Quality of Life lab: from Behavior Change to Quality of Life

Vlad Manea, Katarzyna Wac. *mQoL: Mobile Quality of Life lab: from Behavior Change to Quality of Life*. Workshop on Mobile Human Contributions (MHC 2018), Conference on Pervasive and Ubiquitous Computing (UbiComp 2018). 6p. DOI: `https://doi.org/10.1145/3267305.3267549`. [33]

Abstract Nowadays, the app stores host a variety of mobile health solutions. Smartphone users can choose from tens of thousands of applications, designed to prevent or manage certain diseases, or induce behavior change to improve health and life quality in general. However, the value of most applications remains unclear, as they stop short from documenting adherence to medical evidence. We review the fundamental mobile health challenges and propose Mobile Quality of Life Lab (mQoL), a mobile health platform which addresses the identified challenges and leverages recent developments to facilitate the deployment of much-needed longitudinal, multidimensional, evidence-based studies that are minimally obtrusive for the participants, yet provide high value in terms of the collected datasets, as well as potential for behavior change towards improving Quality of Life.

Keywords mobile application, longitudinal data, behavioral marker, self-assessment, quality of life.

mQoL: Mobile Quality of Life Lab: From Behavior Change to QoL

Vlad Manea
University of Copenhagen
Copenhagen, Denmark
manea@di.ku.dk

Katarzyna Wac
University of Geneva
Geneva, Switzerland
University of Copenhagen
Copenhagen, Denmark
katarzyna.wac@unige.ch
wac@di.ku.dk

UbiComp/ISWC '18 Adjunct, October 8–12, 2018, Singapore, Singapore

ACM ISBN 978-1-4503-5966-5/18/10...$15.00.
https://doi.org/10.1145/3267305.3267549

Abstract
Nowadays, the app stores host a variety of mobile health solutions. Smartphone users can choose from tens of thousands of applications, designed to prevent or manage certain diseases, or induce behavior change to improve health and life quality in general. However, the value of most applications remains unclear, as they stop short from documenting adherence to medical evidence. We review the fundamental mobile health challenges and propose Mobile Quality of Life Lab (mQoL), a mobile health platform which addresses the identified challenges and leverages recent developments to facilitate the deployment of much-needed longitudinal, multidimensional, evidence-based studies that are minimally obtrusive for the participants, yet provide high value in terms of the collected datasets, as well as potential for behavior change towards improving Quality of Life.

Author Keywords
Mobile Application; Longitudinal Data; Behavioral Marker; Self-Assessment; Quality of Life

ACM Classification Keywords
H.4.m [Information systems applications]: Miscellaneous

Introduction
There is a growing need for transdisciplinary efforts towards understanding fundamental theories of Quality of Life (*QoL*)

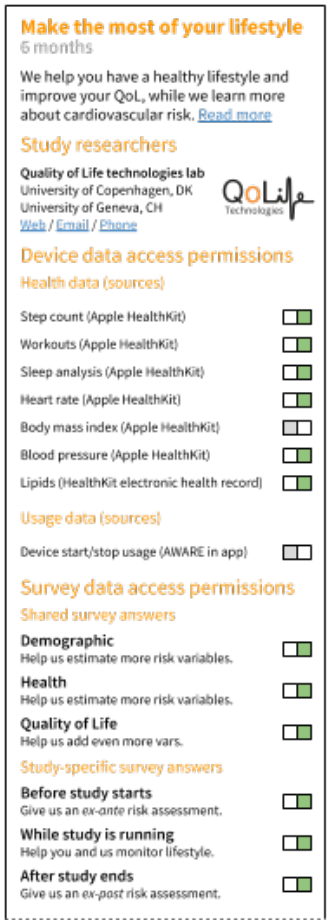


Figure 1: mQoL study standard format: title, duration, description, researchers, data permission requests: device-reported (health and usage) and self-reported (shared and study-specific).

and linking these to an understanding of complex practical problems related to assessing day-to-day individual's QoL [11]. According to the World Health Organization (WHO), the QoL is *"individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns"* [4]. The QoL state is mainly influenced by (un)healthy lifestyle over long periods (*longitudinal*) and multiple dimensions (*multidimensional*) of behavior.

In parallel, personal, miniaturized devices have programmable sensors which are becoming more accurate as technology progresses [3] and collect multiple dimensions of data simultaneously, continuously, in time, and in context. Enabled by them, mobile health (*mHealth*) apps have become the artifact of choice for many recent exploratory and behavior analytics and change studies, conducted by researchers and companies alike. Within mHealth, participants engage in data collection and interventions through numerous channels (e.g., text, audio, graphics, video [14]) and generate device-collected behavioral *markers* and self-reports.

Interested individuals look for mHealth apps in the same stores as any other apps [14], but as of today, there are over 43 thousand mHealth apps [10]. Ideally, the apps would enable effective behavior change towards QoL improvement in the long-term. However, the actual value added by most is unclear due to a lack of medical evidence. Additionally, for researchers, mHealth studies are challenging to conduct, as there is no open, versatile platform enabling the deployment of longitudinal, multidimensional, and evidence-based studies.

mHealth challenges

SCIENTIFIC RIGOR: Poor scientific rigor characterizes most mHealth apps. They do not identify, apply, or document behavior change theories and techniques [8]. Their studies lack bias assessment in participant groups [1], miss control groups [17], or contaminate them with (access to) interventions [1]. They make recommendations without following evidence-based medical guidelines and best practices [9] and are not anchored in a real-life context. Regulation was attempted by international bodies (FDA in the US, TGA in Australia, MDD in Europe) and research initiatives [10], but no definitive standard emerged, putting a burden of choosing useful and harmless mHealth apps on the user.

HOLISTIC ASSESSMENT: In general, apps considering individual participant characteristics, such as wellbeing, lifestyle, personality, and changing needs, can facilitate a holistic view [14], improve the effectiveness of interventions [12], and keep participants engaged longer. Many apps, usually commercial, focus on overall lifestyle, health, and wellbeing, but with unclear effects of behavior change interventions. Other research apps contain healthy behavior change interventions, e.g., combatting sedentarism or quitting smoking, yet they focus only on preventing or managing specific diseases, e.g., diabetes or dementia.

DATA DIMENSIONALITY: For behavior change, feedback based on multidimensional data (e.g., physical and psychological state) yields stronger motivation and avoids reporting flaws [16]. In some studies, multidimensional data is necessary. However, few studies address multiple dimensions [5] and few datasets integrate device data with other data types [6], e.g., blood tests, not only because obtaining the latter is difficult, but also because many researchers choose self-reports instead of reliable data sources [9].

DATA TIMESPAN: mHealth enables the gathering of longitudinal data, allowing the observation of short, medium, and long-term effects. However, many studies continue to focus on short-term data acquisition [5] and involve small sam-

Figure 2: mQoL study standard format (continued): scientific evidence (behavior change and medical), participation consent, and participant signature.

ples of participants, ranging in the tens. Additionally, only a few apps (e.g., [15]) are kept novel for prolonged periods, leading to diminishing effects, interruptions in data collection, and attrition [1]. If larger samples were recruited, then studies would continue to report small improvements, but impactful over the whole population [5].

DATA CONTROL: mHealth apps provide the opportunity for researchers to access and create massive datasets [9]. Apps need to manage these datasets securely and provide complete and accurate information about data generation, measurement, collection, retrieval, and analysis. However, many apps even lack an adequate privacy policy [10].

MHEALTH BURDEN: Researchers are forced to treat the mHealth app as only one aspect of the study, making it difficult to satisfy modern participant expectations regarding maintenance, support, and updates [1, 5, 9] or implement behavior change features (e.g., personalized messages, reminders, or dashboards [13]). Instead, the harsh reality is that researchers often need to keep the app alive in between rounds of funding.

To our knowledge, there is no holistic mobile app for researchers and smartphone users to deploy and participate in evidence-based longitudinal, multidimensional studies to change behaviors and improve QoL in the long-term.

mQoL solution

Faced with these challenges ourselves and aiming at holistic QoL assessment based on behavior change interventions in our QoL technologies lab, we are researching a *Mobile Quality of Life Lab* platform, denoted *mQoL* and operationalized via a mHealth app in Apple iOS, bringing the following benefits for researchers and participants.

RESEARCHERS can now only focus on designing the studies. They can obtain rich behavioral datasets by retrieving longitudinal, multidimensional behavioral markers in time and context, as well as self-reported demographic, medical, and QoL information from participants, all consented, pseudonymized, and structured. The platform is designed to accommodate only exploratory and interventional studies grounded in medical evidence. Its components are designed to maximize participant retention while minimizing study participation burden.

PARTICIPANTS can make sense of behavior and life quality and potentially change behaviors in the long-term, by using only evidence-based studies. While participating, they receive personalized, timely, and contextual information from studies, helping them monitor, observe, and reflect upon daily life and its long-term health and QoL consequences.

mQoL architectural choices

STUDIES: The central concept of mQoL is the *study*, that acts as a research template. Within a study, researchers specify motivations and expectations, provide scientific evidence, plan interventions, specify the needed types of device-reported data, and schedule the retrieval of self-reported data. All studies follow a standard format. Figures 1 and 2 depict a study in this format, on behavior change for reducing cardiovascular disease risk, a current topic in our research. mQoL also allows parallel tracks within the same study, enabling, e.g., both control and intervention groups to participate. We call studies that do not require active participation, e.g., self-reports, *silent*. To avoid contamination of control and intervention groups, mQoL allows at most one non-silent study participation at a time.

DEVICE-REPORTED DATA: mQoL uses device-reported individual health data and smartphone usage data. Health-related data is generated, measured, and collected on the

mQoL surveys

QoL survey: provides a basic, yet holistic view of the participant's QoL. It is based on, e.g., the WHOQOL-BREF [4] validated scale and ideally collected every two weeks.

Demographic survey: helps researchers recruit samples of interest for studies. It contains questions about, e.g., the age, gender, and country, and it is collected rarely.

Medical survey: helps researchers recruit samples of interest, personalize health-related behavior change interventions, and avoid silly recommendations. It includes questions about medication, participant diseases, and family history, and it is collected infrequently.

Personal survey: allows participants to provide contact information, e.g., email or phone, to receive updates about new studies on those channels.

device, continuously, unobtrusively, and independently of the app. All such data can then be retrieved by each study upon consent. Additionally, mQoL allows researchers to design custom tasks in studies, e.g., asking the participant to perform a short-term activity such as a six-minute walk test, from which device-reported data is collected.

SELF-REPORTED DATA: mQoL allows studies to design and schedule self-reported surveys and request access to any of the following self-reported surveys, *shared* between studies: a *Quality of Life* survey, a *demographic* survey, and a *health* survey. An additional *personal* self-reported survey is only visible to mQoL providers. For details, see the **mQoL surveys** side note.

MODULES: mQoL for participants is organized in five tabs, each tab having its modules (Figure 3). (1) The *Data* tab contains modules for managing retrieved data: which studies retrieve which data, and options to pause, restore, stop, and delete each type of data within each study. (2) The *Explore* tab contains two lists with the active and available studies. From this tab, participants can see information (e.g., dashboards) in active studies and can signup for an available study. When the number of studies increases in our app, we plan to design an onboarding feature to help participants choose those that suit their interests and can benefit them most. (3) The *Account* tab contains modules for managing the token and for answering all shared self-reported surveys. (4) The *Settings* tab contains the privacy policy, terms and conditions, and other minor functionalities, e.g., notification management. (5) The *Control* tab is the main entry point in the app. It contains transient, inversely chronological *cards*, which provide information or require action inside modules of the other tabs. Some cards can be announced by notifications with reminders and personalized messages. See a clickable mock-up at http://bit.ly/mobileQoLlab.

TECHNOLOGIES: mQoL leverages the Apple iOS platform, for several reasons. First, the App Store has a stricter review process than other platforms, yielding to more qualitative apps. Then, iOS allows device-reported and self-reported data collection via often-used and well-documented frameworks, making study design and participation experience familiar. Last, Apple continues to invest in digital health at scale (e.g., they released an API for electronic health records in June 2018). For details, see the **mQoL technologies** side note.

mHealth challenges vs. mQoL

SCIENTIFIC RIGOR: By reviewing the mandatory scientific sources included in the studies and their implementation, including requiring an external review, e.g., an ethical approval, mQoL ensures all studies rely on the latest medical evidence. While this model is strict and laborious, it helps mQoL become the authoritative app for scientific studies researchers and participants ultimately need.

PARTICIPANT ASSESSMENT; DATA DIMENSIONALITY AND TIMESPAN: mQoL addresses these challenges through the retrieval of continuous device-reported and scheduled device- and self-reported data, performed in parallel and over long periods as part of studies.

DATA CONTROL: For pseudonymous data retrieval, upon installing mQoL, the participant sets up a *token*. This token (and no other personal information) will identify the data retrieved from the participant. Such an approach has been used in recent health studies [7]. For retrieving data, each study requests only the most granular data types it needs, e.g., physical activity $\rightarrow$ walking $\rightarrow$ steps $\rightarrow$ daily count. However, for studies which need to transfer data out of the device, the app securely transmits data upon separate con-

mQoL technologies

Apple HealthKit: framework used to collect and retrieve individual health data, including electronic health records (since June 2018). https://developer.apple.com/healthkit/

AWARE: framework used to collect and retrieve smartphone usage data. http://www.awareframework.com/

Apple ResearchKit: framework used to design consents, surveys, and tasks for the participant. https://developer.apple.com/researchkit/

Charts: library used to draw interactive dashboards. https://github.com/danielgindi/Charts

Open mHealth: schemas used as a format for health data exported outside of the device. http://www.openmhealth.org/

Parse: library used to export data to the mQoL Smart Lab [2], which uses this technology. https://docs.parseplatform.org/ios/guide/

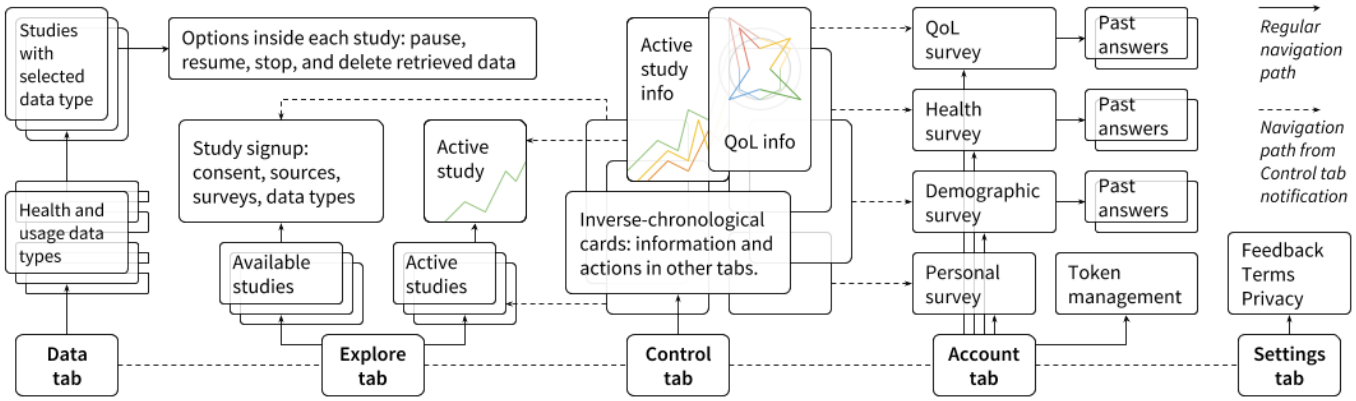


Figure 3: Conceptual architecture of *mQoL*: tabs and modules.

sent. To retrieve or export data, mQoL needs to send the participant a notification, which implies a permission request for every processing.

MHEALTH BURDEN: mQoL simplifies study deployment for researchers, by providing a platform that provides a format for designing studies as well as well-documented and often-used frameworks and libraries for consents, tasks, health data, usage data, and survey data, helping them worry less about app maintenance or study survival.

Conclusion and further work

We reviewed the benefits, needs, and shortcomings of the mHealth domain and observed that there is no holistic platform for researchers and smartphone users that allows them to conduct and participate in evidence-based longitudinal, multidimensional studies. We propose *mQoL*, a mobile platform designed to addresses this gap as well as the ardent needs of mHealth in general, with the potential of being leveraged in numerous evidence-based studies, to change behaviors and improve QoL. The research is ongoing, and at the moment we are looking into ways of streamlining the study designs for researchers, as well as putting in place mechanisms to evaluate evidence basis for studies. However, it is real-world studies that can ultimately validate mQoL. The first study is our project on behavior change for reducing cardiovascular disease risk, to be deployed later this year. Medical experts engaged in the project will provide feedback on the mQoL platform and study designs.

Acknowledgments

EU H2020 WellCo project (769765): http://wellco-project.eu

REFERENCES

1. O Byambasuren, S Sanders, E Beller, and P Glasziou. 2018. Prescribable mHealth apps identified from an overview of systematic reviews. *npj Digital Medicine* 1, 1 (2018), 12.

2. A De Masi, M Ciman, M Gustarini, and K Wac. 2016. mQoL smart lab: quality of life living lab for interdisciplinary experiments. In *UBICOMP Adjunct*. ACM, 635–640.

3. KR Evenson, MM Goto, and RD Furberg. 2015. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act* 12, 1 (2015), 159.

4. WHOQOL Group and others. 1995. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 41, 10 (1995), 1403–1409.

5. H Klimis, J Thakkar, and CK Chow. 2018. Breaking Barriers: Mobile Health Interventions for Cardiovascular Disease. *Can J Cardiol* (2018).

6. W Lim and others. 2018. Beyond fitness tracking: The use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. *PLoS Biol* 16, 2 (2018).

7. MV McConnell, A Shcherbina, A Pavlovic, JR Homburger, RL Goldfeder, D Waggot, MK Cho, ME Rosenberger, WL Haskell, J Myers, and others. 2017. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts Cardiovascular Health Study. *JAMA Cardiol* 2, 1 (2017), 67–76.

8. S Michie, M Richardson, M Johnston, C Abraham, J Francis, W Hardeman, MP Eccles, J Cane, and CE Wood. 2013. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 46, 1 (2013), 81–95.

9. AC Moller, G Merchant, DE Conroy, R West, E Hekler, KC Kugler, and S Michie. 2017. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. *J Behav Med* 40, 1 (2017), 85–98.

10. L Neubeck, N Lowres, EJ Benjamin, SB Freedman, G Coorey, and J Redfern. 2015. The mobile revolution - using smartphone apps to prevent cardiovascular disease. *Nat Rev Cardiol* 12, 6 (2015), 350.

11. M Nussbaum and A Sen. 1993. *The quality of life*. Oxford University Press.

12. L Piwek, DA Ellis, S Andrews, and A Joinson. 2016. The rise of consumer health wearables: promises and barriers. *PLoS Med* 13, 2 (2016).

13. R2G. 2016. mHealth Economics 2016 - Current Status and Trends of the mHealth App Market. (2016).

14. CMS Sarwar and others. 2018. Mobile health applications in cardiovascular research. *Int J Cardiol* (2018).

15. GH Tison, K Hsu, JT Hsieh, BM Ballinger, MJ Pletcher, GM Marcus, and JE Olgin. 2017. Achieving High Retention in Mobile Health Research Using Design Principles Adopted From Widely Popular Consumer Mobile Apps. (2017).

16. MJ Western, OJ Peacock, A Stathi, and D Thompson. 2015. The understanding and interpretation of innovative technology-enabled multidimensional physical activity feedback in patients at risk of future chronic disease. *PloS One* 10, 5 (2015).

17. SJ Winter, JL Sheats, and AC King. 2016. The use of behavior change techniques and theory in technologies for cardiovascular disease prevention and treatment in adults: a comprehensive review. *Prog Cardiovasc Dis* 58, 6 (2016), 605–612.

Publication 3: Call for Papers for the Workshop on Longitudinal Data Collection in Human Subject Studies

Vlad Manea, Allan Berrocal, Alexandre De Masi, Naja Holten Møller, Katarzyna Wac, Hannah Bayer, Sune Lehmann, Euan Ashley. *Call for Papers: LDC '19: Workshop on Longitudinal Data Collection in Human Subject Studies*. Call for Papers for the Workshop on Longitudinal Data Collection in Human Subject Studies (LDC 2019), Conference on Pervasive and Ubiquitous Computing (UbiComp 2019). 4p. DOI: `https://doi.org/10.1145/3341162.3347758`. [20]

Abstract Individuals increasingly use mobile, wearable, and ubiquitous devices capable of unobtrusive collection of vast amounts of scientifically rich personal data over long periods (months to years), and in the context of their daily life. However, numerous human and technological factors challenge longitudinal data collection, often limiting research studies to very short data collection periods (days to weeks), spawning recruitment biases, and affecting participant retention over time. This workshop is designed to bring together researchers involved in longitudinal data collection studies to foster an insightful exchange of ideas, experiences, and discoveries to improve the studies' reliability, validity, and perceived meaning of longitudinal mobile, wearable, and ubiquitous data collection for the participants.

Keywords longitudinal study, human-subject study, human sensing, mobile device, in-situ, panel technique, attrition.

LDC '19: International Workshop on Longitudinal Data Collection in Human Subject Studies

Vlad Manea
University of Copenhagen (DK)
manea@di.ku.dk

Allan Berrocal
University of Geneva (CH)
allan.berrocal@unige.ch

Alexandre De Masi
University of Geneva (DK)
alexandre.demasi@unige.ch

Naja Holten Møller
University of Copenhagen (DK)
naja@di.ku.dk

Katarzyna Wac
University of Copenhagen (DK)
University of Geneva (CH)
wac@di.ku.dk

Hannah Bayer
Datacubed Health (USA)
hannah@datacubed.com

Sune Lehmann
Technical University of Denmark (DK)
sljo@dtu.dk

Euan Ashley
Stanford University (USA)
euan@stanford.edu

ABSTRACT

Individuals increasingly use mobile, wearable, and ubiquitous devices capable of unobtrusive collection of vast amounts of scientifically rich personal data over long periods (months to years), and in the context of their daily life. However, numerous human and technological factors challenge longitudinal data collection, often limiting research studies to very short data collection periods (days to weeks), spawning recruitment biases, and affecting participant retention over time. This workshop is designed to bring together researchers involved in longitudinal data collection studies to foster an insightful exchange of ideas, experiences, and discoveries to improve the studies' reliability, validity, and perceived meaning of longitudinal mobile, wearable, and ubiquitous data collection for the participants.

CCS CONCEPTS

• **Human-centered computing** → *Ubiquitous and mobile computing theory, concepts and paradigms*; **Ubiquitous and mobile computing design and evaluation methods**; Empirical studies in ubiquitous and mobile computing.

KEYWORDS

Longitudinal studies, Human subject studies, Human sensing, Mobile devices, In situ, Panel technique, Attrition

ACM Reference Format:
Vlad Manea, Allan Berrocal, Alexandre De Masi, Naja Holten Møller, Katarzyna Wac, Hannah Bayer, Sune Lehmann, and Euan Ashley. 2019. LDC '19: International Workshop on Longitudinal Data Collection in Human Subject Studies. In *Proceedings of UbiComp/ISWC Adjunct '19: Adjunct Proceedings of the International Conference on Ubiquitous Computing (UbiComp/ISWC Adjunct '19).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/1122445.1122456

UbiComp/ISWC Adjunct '19, September 11–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 MOTIVATION

There is a need for longitudinal data collection (LDC) in human subject studies for accurate observational and intervention purposes, especially in the health domain, as the long-term repetitive behaviors and lifestyle choices influence individuals' health outcomes and quality of life (QoL) in the long term. However, oftentimes, data collected over long periods mainly relies on repeated, but momentary self-reports. Mobile, wearable, and ubiquitous (MWU) devices can unobtrusively collect continuous data that provides numerous benefits unavailable otherwise, e.g., help to uncover gradual changes in behaviors related to aging [6], assess the risk of chronic diseases [2], support the disease diagnosis [22], and understand medication compliance [8], all of which add value to optimize interventions, increase longevity, and improve QoL in the long term.

2 LONGITUDINAL MWU STUDY FACTORS

Numerous factors affect the quantity and quality of MWU data collected during long-term research studies, particularly in the areas of wellness, health, and quality of life.

Participants' attitudes towards self-monitoring, overall goals toward health and fitness, values and desires at a given time, and concerns over data security are only a few **human factors** in the participants' recruitment phase, as well as data collection [15]. Other factors include mismatches between the service expectations and user experience, the discomfort with the presented information, and the effort needed to use the MWU device in daily life [4, 14]. Long-term compliance follows from device usage behaviors (e.g., weekdays vs. weekends, working days vs. holidays) and the purpose of behavior monitoring (e.g., daily steps for sedentary individuals vs. fitness levels for achievers) [13].

Technical factors include high data dimensionality, heterogeneity, temporal dependency, sparsity, irregularity, noisiness, ambiguity, and redundancy [15]. Potential interaction-related reasons for study abandonment include potential platform issues, e.g., tracking

accuracy, usability issues, synchronization problems, and battery life [14].

These factors pose significant challenges not only on the potential study participants, providing their datasets and striving for minimal obtrusiveness and maximum gain (including monetary incentives), but also the study owners, aiming to maximize data quality and results' generalization.

3 LONGITUDINAL MWU STUDIES

Longitudinal studies were deployed in the past. From the perspective of the number of participants, at the lower end lie studies performed at repeated intervals in independent labs involving tens of participants by using, for the time of the lab study, single-purpose MWU artifacts that collect data for specific behaviors (e.g., physical activity on a treadmill, sleep with an oximeter for one night). At the higher end, longitudinal studies involving hundreds of thousands of people across generations and decades traditionally only contain data provided by the clinician into healthcare systems, e.g., during visits [20] and self-reported data provided by participants via unassisted questionnaires, e.g., mailed regularly [3]. Recently, some studies included MWU data (Table 1). However, due to the high effort and price of LDC at scale, they remain limited to a few days or months of data. Large-scale longitudinal human subject studies routinely leveraging MWU data are rare, and some are ongoing (Table 2).

Improving the quality of the results is essential for the scientific community. Longitudinal study designs with MWU data can benefit by unobtrusively collecting accurate data from daily life, from representative samples of participants, over long periods, where behaviors unfold. However, such studies are difficult in practice.

Original Study	Sub-Study Involving MWU Data
Framingham Heart Study	N=790 participants from the original cohorts of the Framingham Heart Study monitored blood pressure and heart rate for 12 weeks with the Nokia-Withings BP-801 blood pressure cuff and the Apple Watch Generation 0 smartwatch. 44% of those who received at least one device provided data in the last week [18].
Framingham Health Study	N=830 participants from the original cohorts of the Framingham Heart Study participants monitored sleep for at most two consecutive nights with MyCardio and Nonin devices for for electrocardiogram and oximetry, respectively [11].
National Health and Nutrition Examination and Survey	N=11.959 participants (4.028 youth and 7.931 adults) monitored physical activity for 8 hours (minimum) in 3 days (minimum) with the ActiGraph AM-7164 accelerometer [10].
Nurses' Health Studies	N=121.700 (Period I, since 1976), N=116.678 (Period II, since 1989), N=280.000+ (Period III, since 2010) shared wearable physical activity and sleep (7 days, 4 times/year) [5].
UK Biobank	N=103.712 participants (44.8% response) monitored physical activity for 6.9 days on average with the Axivity AX3 accelerometer [9].
Women's Health Study	N=16.741 participants monitored physical activity for 10 hours (minimum) in 4 days (minimum) with ActiGraph GT3X+ accelerometer. [19]

Table 1: Longitudinal Studies with Rudimentary MWU Data

4 WORKSHOP OBJECTIVES

The workshop provides professional space for researchers to share ideas, approaches, methods, tools, frameworks, and other insights that enable the collection of reliable and valid longitudinal MWU data. The workshop aims to present and discuss state of the art methods for longitudinal MWU data collection in human subject studies. We aim to discuss ideas to minimize participants' burden while maximizing their retention. Such studies contribute relevant data to support the replicability of the study results that will create value for researchers and participants alike. Then, we aim at mapping the challenges into the implications for the design of human subject studies that will drive this line of research in the coming years. Finally, we aim to foster collaboration among researchers working in this area.

5 WORKSHOP CONTRIBUTIONS

We accepted three papers for publication and presentation at the workshop. Vasconcelos et al. [24] discuss the challenges and lessons learned in four longitudinal studies with older adults and chronic disease patients in the context of assessment of self-care technologies. Majethia et al. [16] design a long term data-driven study on a finite student population of a residential university campus to gain a comprehensive understanding of how groups form and evolve over time. Van Berkel et al. [23] propose the use of game theory in longitudinal mobile sensing deployments towards capturing contextual morality while keeping a high level of engagement.

The papers published in this workshop contribute in, but not limited to, the following areas: elaboration on human, technological, and other significant factors influencing the design and execution of the LDC in human subject studies [23, 24]; approaches that are likely to increase the quality of MWU data collected as part of scientific studies or identify participant groups likely to exhibit high compliance [24]; methodologies to assess and improve retention for a representative sample of participants and specific metrics, e.g., engagement, interruptions, consistency, or time to abandonment [23]; novel findings and lessons learned from past or existing

Study	Description
AllOfUs	N=1.000.000+ US residents. Behavioral and medical outcomes assessment. In recruitment.
Heart eHealth Study	N=1.000+, cardiovascular health, integrates with 9 wearable and mobile data providers.
Kavli Human Project	N=10.000 New Yorkers from 2500 households for 20 years. In recruitment.
mQoL Living Lab	N=1.000+, integrates with 4 wearable and mobile data providers. Multiple studies.
MyHeart Counts	N=48.900+, cardiovascular health, mobile application collecting health data (7 days).
Open Humans	N=4.800+, integrates with 20+ genetic, wearable, and mobile data providers. Multiple studies.
Project Baseline	N=10.000 US residents for 5 years, integrates mobile and wearable data. In recruitment.
Social Fabric Project	N=1.000 Danish students given a mobile device that logged their social interactions.

Table 2: Ongoing Longitudinal Studies with MWU Data

LDC studies conducted in the user's context, implying qualitative, quantitative, or mixed analyses [24]; and techniques or methods to analyze the representatives and quality of collected longitudinal MWU data [16].

6 WORKSHOP SUBMISSION

The accepted papers were submitted in two iterations. In the first iteration, each article received reviews from three committee members, who collectively decided to accept them conditionally. In the second iteration, the authors argued for their paper and submitted revised versions. While there are only three papers accepted, we welcome all interested conference attendees to the workshop and expect 10 to 20 participants in total.

7 WORKSHOP SCHEDULE

The proposed workshop on Longitudinal Data Collection is planned to cover a half day. Following a round of introduction by all attendees, we expect a keynote speech by a senior researcher with extensive experience in conducting longitudinal studies. Following the keynote, the organizers will introduce the papers, after which the respective authors will present their work for 10-20 minutes.

Time	Activity
09 : 15 – 09 : 30	Opening Notes
09 : 30 – 10 : 00	Keynote
10 : 00 – 11 : 00	Presentations
11 : 00 – 11 : 15	Coffee Break
11 : 15 – 12 : 15	Brainstorm Sessions
12 : 15 – 12 : 30	Lessons Learned
12 : 30 – 12 : 35	Closing
18 : 00 – …	Joint Dinner

Table 3: Workshop Schedule

Following the presentations, all attendees will contribute to a workshop and discussion to identify the implications for the broader research agenda. An organizer (KW) will summarize the results of the ongoing discussions and present them to the audience in a session of lessons learned elicitation. We expect these findings to fuel dialogue over dinner and spark future collaborations among attendees. The workshop will have a break aimed at incentivizing attendees to discuss between themselves. Finally, attendees are invited to a joint dinner. The schedule is depicted in Table 3.

8 WORKSHOP ORGANIZERS

The organizers gained vast experience in LDC studies. Prof. Wac is the principal investigator (PI) in the mQoL Living Lab [7], Alexandre, Allan, and Vlad conduct LDC studies in the same lab. Prof. Holten Møller researches qualitative methods assuring quantitative data quality collected in situ, especially in healthcare settings [12]. Prof. Bayer is the chief scientist in The Kavli Human Project [1], Prof. Ashley is the PI of MyHeart Counts [17], and Prof. Lehmann was the PI for Social Fabric Project [21]. All organizers learned hard lessons on longitudinal study design and execution, and study participant retention, also related to, amongst the others, the MWU choices, above and beyond the protocols' timeframes.

Vlad Manea is Ph.D. student in computer science at the University of Copenhagen, Denmark. His research interests include ubiquitous computing, mobile health, and machine learning.

Allan Berrocal is Ph.D. student in computer science and Swiss Government excellence scholar at the University of Geneva, Switzerland. His research interests include human-computer interaction, pervasive and mobile computing, mobile health, and human stress.

Alexandre DeMasi is Ph.D. student in computer science at the University of Geneva, Switzerland. His research interests include pervasive and mobile computing, quality of experience, context awareness, and machine learning.

Naja Holten Møller, Ph.D. is assistant professor of computer science at the University of Copenhagen, Denmark. Her research interests include computer-supported cooperative work, human-computer interaction, science and technology studies, ethnography, and workplace studies.

Katarzyna Wac, Ph.D. is associate professor of computer science and leader of the Quality of Life Technologies Lab at the University

of Copenhagen, Denmark and the University of Geneva, Switzerland, affiliated with Stanford University. Her research interests include pervasive and mobile computing, behavior modeling, digital health, quality of experience, and quality of life.

Hannah Bayer, Ph.D. is the chief scientific officer at Datacube and former research associate professor at the New York University, United States. Her research interests include decision making, human condition, big data, and urban studies.

Sune Lehmann, Ph.D. is professor of computer science at the Technical University of Denmark. His research interests include complex networks, social networks, social data.

Euan Ashley, Ph.D. is professor of cardiovascular medicine at Stanford University, United States. His research interests include genomics, precision medicine, inherited cardiovascular disease, personalized medicine, and cardiomyopathy.

9 SUMMARY

The Workshop on Longitudinal Data Collection foregrounds contributions and facilitates discussions focused on the methods, tools, and frameworks for collection, analysis, and interpretation of human subjects' data obtained over long periods. This workshop is not only a valuable but also a timely and relevant addition to the UbiComp conference and the community at large.

ACKNOWLEDGMENTS

WellCo EU H2020 Project No. 769765. ECoWeB EU H2020 Project No. 754657. MIQmodel SNF Project No. 157003.

REFERENCES

[1] O Azmak et al. 2015. Using big data to understand the human condition: the Kavli HUMAN project. *Big Data* 3, 3 (2015), 173–188.
[2] FW Booth et al. 2011. Lack of exercise is a major cause of chronic diseases. *Comprehensive Physiology* 2, 2 (2011), 1143–1211.
[3] K Breckenridge et al. 2015. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. *Nephrology Dialysis Transplantation* 30, 10 (2015), 1605–1614.
[4] AI Canhoto et al. 2017. Exploring the factors that support adoption and sustained use of health and fitness wearables. *Journal of Marketing Management* 33, 1-2 (2017), 32–60.
[5] GA Colditz and SE Hankinson. 2005. The Nurses' Health Study: lifestyle and health among women. *Nature Reviews Cancer* 5, 5 (2005), 388.
[6] PN Dawadi et al. 2013. Automated assessment of cognitive health using smart home technologies. *Technology and Health Care* 21, 4 (2013), 323–343.
[7] A De Masi et al. 2016. mQoL smart lab: quality of life living lab for interdisciplinary experiments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. Association for Computing Machinery, 635–640.
[8] A de Thurah et al. 2010. Compliance with methotrexate treatment in patients with rheumatoid arthritis: influence of patients' beliefs about the medicine. A prospective cohort study. *Rheumatology International* 30, 11 (2010), 1441–1448.
[9] A Doherty et al. 2017. Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study. *PloS one* 12, 2 (2017), e0169649.
[10] KR Evenson and F Wen. 2015. Performance of the ActiGraph accelerometer using a national population-based sample of youth and adults. *BMC Research Notes* 8, 1 (2015), 7.
[11] EJ Heckman et al. 2017. Wearable sleep epidemiology in the Framingham Heart Study. *Journal of Sleep and Sleep Disorders Research* 40, suppl_1 (2017), A289–A289.
[12] NL Holten Møller et al. 2017. Data tracking in search of workflows. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Association for Computing Machinery, 2153–2165.
[13] S Ickin et al. 2012. Factors influencing quality of experience of commonly used mobile applications. *IEEE Communications Magazine* 50, 4 (2012), 48–56.
[14] H Jeong et al. 2017. Smartwatch wearing behavior analysis: a longitudinal study. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 60.
[15] J Kim et al. 2018. Enchanted life space: adding value to smart health by integrating human desires. *J Healthc Inform Res* 24, 1 (2018), 3–11.
[16] R Majethia et al. 2019. Cohort analyses of in-person interactions in temporally evolving student social groups. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. Association for Computing Machinery.
[17] MV McConnell et al. 2017. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts Cardiovascular Health Study. *JAMA Cardiology* 2, 1 (2017), 67–76.
[18] DD McManus et al. 2019. Design and Preliminary Findings From a New Electronic Cohort Embedded in the Framingham Heart Study. *Journal of Medical Internet Research* 21, 3 (2019), e12143.
[19] P Phongsavan et al. 2004. Estimating physical activity level: the role of domestic activities. *Journal of Epidemiology & Community Health* 58, 6 (2004), 466–467.
[20] GA Roth. 2018. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 392, 10159 (2018), 1736–1788.
[21] A Stopczynski et al. 2014. Measuring large-scale social networks with high resolution. *PloS one* 9, 4 (2014), e95978.
[22] EJ Topol. 2019. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine* 25, 1 (2019), 44.
[23] N Van Berkel et al. 2019. Capturing contextual morality: applying game theory on smartphones. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. Association for Computing Machinery.
[24] A Vasconcelos et al. 2019. Challenges and lessons learned from implementing longitudinal studies for self-care technology assessment. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. Association for Computing Machinery.

Publication 4: mQoL Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices

Allan Berrocal, Vlad Manea, Alexandre De Masi, Katarzyna Wac. *mQoL-Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices*. Conference on Mobile Systems and Pervasive Computing (MobiSPC 2020). 9p. DOI: `https://doi.org/10.1016/j.procs.2020.07.033`. **Nominated for the best project pitch award at the University Hospitals of Geneva Innovation Day 2020, Geneva, Switzerland**. [21]

Abstract Human subject studies with mobile users are widely used to understand, and model, human aspects such as behaviours and preferences, in the lab and in the wild. These studies usually employ mixed methods, collecting data by active participation and passive sensing using interactive, mobile, wearable, and ubiquitous devices. Researchers rely on a software platform to design and execute their studies, but existing solutions require a steep learning curve, allow little control, and offer limited guarantees. Our research lab built the mQoL-Lab platform using open source technologies, and evolved it to a durable and reliable software ecosystem in over ten mobile subject studies along eight years across three countries. In this paper, we share the acquired experience via tangible artifacts such as requirements, architecture, design, step-by-step support, configuration scripts, and recommendations for researchers to construct a software platform supporting mobile subject studies. The paper is especially relevant for researchers embracing short-term to longitudinal, observational or intervention-based studies, leveraging mixed methods, including multiple devices, and tens to hundreds of simultaneous participants.

Keywords mobile study, mobile platform, mixed methods, passive sensing, mobile interaction, wearable devices, data collection.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 175 (2020) 221–229

Procedia Computer Science

www.elsevier.com/locate/procedia

The 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC)
August 9-12, 2020, Leuven, Belgium

mQoL Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices

Allan Berrocal[a], Vlad Manea[b], Alexandre De Masi[a], Katarzyna Wac[a,b,*]

[a] *Quality of Life Technologies Lab, Institute of Service Science, University of Geneva, Switzerland*
[b] *Quality of Life Technologies Lab, Department of Computer Science, University of Copenhagen, Denmark*

Abstract

Human subject studies with mobile users are widely used to understand, and model, human aspects such as behaviours and preferences, in the lab and in the wild. These studies usually employ mixed methods, collecting data by active participation and passive sensing using interactive, mobile, wearable, and ubiquitous devices. Researchers rely on a software platform to design and execute their studies, but existing solutions require a steep learning curve, allow little control, and offer limited guarantees. Our research lab built the *mQoL Lab* platform using open source technologies, and evolved it to a durable and reliable software ecosystem in over ten mobile subject studies along eight years across three countries. In this paper, we share the acquired experience via tangible artifacts such as requirements, architecture, design, step-by-step support, configuration scripts, and recommendations for researchers to construct a software platform supporting mobile subject studies. The paper is especially relevant for researchers embracing short-term to longitudinal, observational or intervention-based studies, leveraging mixed methods, including multiple devices, and tens to hundreds of simultaneous participants.

Keywords: mobile studies, mobile platform, mixed methods, passive sensing, mobile interaction, wearable devices, data collection

1. Introduction

Human subject studies with mobile users in the wild, i.e., outside the research lab, are widely used as a method to better understand human behaviours which occur in the context of daily life, or intervene with behaviors, in case of intervention-based studies. These studies usually imply the use of mixed methods, e.g., qualitative self-reported outcomes referred to as "Participant Provided Outcomes" (inspired by "Patient Reported Outcomes" or PROs from the taxonomy of clinical outcomes [30]) leveraging methods such as Ecological Momentary Assessment (EMA) [33], Day Reconstruction Method (DRM) [23], or longer multi-item surveys. The studies also involve quantitative,

* Corresponding author. Tel.: +41-22-379-024
E-mail address: katarzyna.wac@unige.ch

1877-0509
10.1016/j.procs.2020.07.033

technology-reported outcomes (TechROs [30]) from data collected passively [9] by interactive, mobile, wearable, and ubiquitous (IMWU) devices. Typically, researchers utilize a software platform to conduct the mobile subject studies. Some platforms are paid-per-use, open source, or custom-made by their research group. Selecting or developing a platform to conduct studies appropriately is a nontrivial task, especially for research labs with limited access to software engineering resources or expertise.

Existing platforms partially address this shortcoming by providing off-the-shelf products, or stand-alone components, rich in functionality for mobile subject studies, such as smartphone sensor data collection or survey administration. However, they require a steep learning curve which involves the exploration of the feature set, only a fraction of which is often needed, or used in a study. Also, because they are offered as is, there are few guarantees for support and troubleshooting. Additionally, because mobile subject studies collect personal data, they are inherently subject to strict ethical and legal regulations, such as the General Data Protection Regulations (GDPR) in the EU, or the Health Insurance Portability and Accountability Act (HIPAA) in the USA. Research labs require robust and flexible infrastructures to satisfy changing needs on a timely manner. However, without coordinated planning, and a suitable technical infrastructure, labs resort to creating disparate infrastructures which are difficult to maintain and reuse. As a result, attempts to build platforms result in artefacts which fall short from bringing long-term benefits.

Our research lab conducts mobile subject studies collecting behaviour and health-related data from IMWU devices (e.g., smartphones and wearables) to support behaviour assessment. To this end, our lab has incrementally developed the *mQoL Lab* platform [11], a robust and flexible ecosystem based on a combination of open source dependencies and custom components since 2011. We deployed and used instances of our platform to conduct mobile subject studies in parallel, within separate research areas, study populations, mixed methods, technical environments, device heterogeneity, and data security regulations mandated by ethical protocols in Switzerland, Denmark and the USA.

The main contribution of this paper is the description of the *mQoL Lab* platform to conduct mobile subject studies. Over time, we revised the platform to support studies matching our growing research needs. We therefore find it relevant to describe and share its overall architecture with the community. This paper presents the architecture of the platform and provides key insights for designing, developing, maintaining, and evolving a platform by following these guidance. Our lessons learned and guiding advice are relevant for researchers who are preparing to conduct mobile subject studies involving simultaneous participants (tens to hundreds), from a few days to several years, by employing qualitative, quantitative, or mixed research methods, potentially across geographies.

2. Characterizing Human Subject Studies

2.1. Constituents and Data in Human Subject Studies

Human subject studies usually have three constituents: two “actors” - participants and researchers - and the “system”. Fig. 1 shows the study participants on the left hand side, and the researchers conducting the study on the right hand side. The system is depicted in the middle, as it enables the data collection from participants, by using artefacts such as IMWU devices, and data analysis by the researchers. For instance, in human-computer interaction, during mobile subject studies, the collected data pertains to the interaction between the participant and the artefacts in context; in behavioural science, the collected data pertains to the participant behaviours in context, as measured by the artefacts.

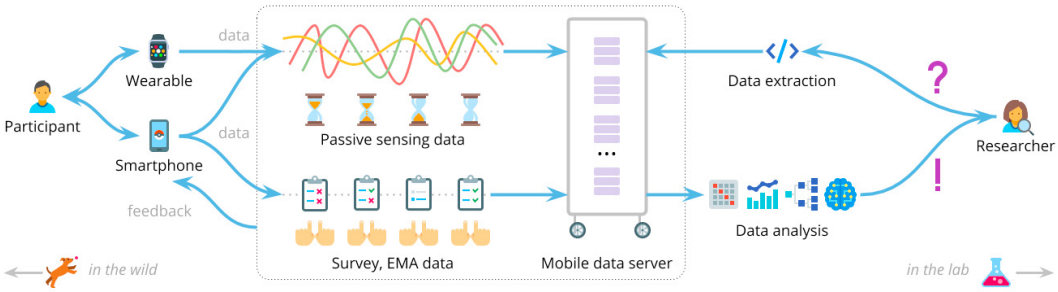


Fig. 1: Overall data flow of a generic human subjects study with mobile users

In mobile subject studies, researchers typically incorporate both qualitative and quantitative data. Qualitative data is usually collected from surveys (whose outcomes rely on scoring of a validated scale). Smartphones can facilitate the collection of data from surveys.

Quantitative data is usually obtained as TechROs collected through passive sensing by IMWU devices (often from the context of daily life). For example, commonly used data from the smartphone itself includes: position and orientation, applications usage, notification events, screen events, network connectivity, ambient light, ambient temperature, battery level, as well as more personal traces such as recognized physical activity, geographical location, ambient sound, calls, messages, audio and video. Data from wearable devices can be of physical (e.g., steps, energy expenditure, distance, and duration of physical activity as well as sleep) and physiological (e.g., electrodermal activity, heart rate, heart rate variability, respiration, glucose levels) nature, as well as other types [35]. After the data collection, researchers typically perform extraction and analysis on both qualitative and quantitative data by following, for example, an iterative hypothetico-deductive approach [28].

2.2. Requirements of Mobile Subject Studies

This section describes common requirements for a research platform to conduct mobile subject studies. These requirements illustrate important functionalities associated with the three constituents identified in Fig. 1: the participants (R1-R4), the system (R5-R8), and the researchers (R9-R11). Instead of using a strict software engineering requirements decomposition, we present them as a combination of functional and non-functional requirements in Table 1. In this context, functional requirements refer to the scope of the system, and stem from study objectives, and researcher investigation experience. Non-functional requirements stem from necessary system properties (e.g., usability aspects for participants and researchers) and researchers' need for gradual automation within the system.

We do not claim that the list of requirements of the *mQoL Lab* is exhaustive. Instead, the requirements emerged over time in the following four life stages of the platform:

Stage 1: Our research began in 2010, in Switzerland, by instrumenting smartphones with a sensor data-logger (acceleration, location, network, screen, applications used, battery state, among others) for brief periods, covering requirements R1, R7, R8 and studies [6, 7, 14, 20, 21, 22].

Stage 2: Then, in 2017, as we were repeating steps for every study in Switzerland and the USA, we evolved the sensors data-logger into a platform that allowed parallel studies with separate configurations, covering requirements R2, R3, R4, R6, R9 and studies [4, 12, 13].

Stage 3: Afterwards, the sensing capacity from smartphones increased, but hardware and software restrictions and policies (from Google and Apple) limited access from development frameworks, which forced updates to the platform. At the same time, adoption and measurement accuracy of wearable devices for daily life outcomes made significant progress. Thus, in 2018 we updated the platform to support a set of consumer-friendly wearable devices, covering updates to requirements R4, R7, R8 and studies [10, 26, 29, 36].

Stage 4: Finally, driven by recent data protection regulations in different countries (specifically, Switzerland, USA and Denmark), in 2019, we enhanced the platform to easily allow re-instantiations, covering requirement R5, R6, and partly R11, with a completed study [2]. Re-instantiations enabled us to deploy the platform to the Stanford University Hospital where we are currently conducting a study for clinical patients collecting longitudinal data from multiple sources: self-reports, peer-reports [2], and technology-reports (mobile application and wearables simultaneously).

2.3. Existing Platforms for Mobile Subject Studies

Research using mobile devices is naturally growing, as smartphones become more ubiquitous. Some researchers created their own mobile applications to record passive data, especially from Android smartphones [34, 1, 27, 8]. Moreover, some research groups developed mobile applications, as well as larger platforms made available for other researchers [16, 27, 24]. Table 2 depicts some of the most popular mobile sensing solutions used in previous research. The systems listed there have slightly different focus areas, but share the common goal to support research in mobile sensing. The first two solutions (AWARE [16] and Sensus [27]) are sufficiently equipped to support mobile subject studies (we label them as mobile frameworks in Table 2). For instance, they work in multiple operating systems (Android and iOS) supporting both passive and active data collection. SensingKit [24] is a specialized library that simplifies the interaction with on-board smartphone sensors. The other solutions in Table 2 are not actively maintained.

Table 1: Requirements and their corresponding motivations for mobile subject studies

	Requirement	Motivation
	Requirements for study participants	
R1	Participants can provide consent and take part in studies at home, over the internet (functional).	Participants should be minimally required to have access to a web browser. For most studies, they are (additionally) required to own an IMWU device, such as their own smartphone or wearable. At times, they receive a device in temporary or permanent ownership. More than 65% of participants prefer being able to contribute to the study at home, over the internet [5].
R2	Participants can control presence in a study (functional).	Participants should receive clear information about the institution, purpose, contributions, and data collected from the study. They should be able to start, pause, resume, stop, and delete their participation from the study at any time, and at no cost. Trust in the research is an important factor for participation in studies [5]. Transparency minimizes concern and confusion while giving informed consent to participate [15].
R3	Participants can control data provision (functional).	Participants should receive clear information about the collection, storage, and analysis of their data as part of the study. They should be able to start, pause, resume, stop, and delete their data at any time and at no cost. For example, participants should be able to answer or skip surveys and authorize/deauthorize IMWU devices. Allowing control and authorization is a determinant factor for individuals to participate in studies [15].
R4	Participants can contribute to studies in the lab and in the wild (functional).	Participants should be able to contribute to a study (either in the wild or in the lab) by providing passive sensing (e.g., using a smartphone or wearable device), active involvement (e.g., answering surveys), or a combination of both, depending on the measured outcomes of the study [31] defined by the researcher.
	Requirements for study researchers	
R5	The mobile data server can be easily re-instantiated (non-functional).	The mobile data server should allow deployment on Linux-based host environments with minimal *a priori* dependencies. Components should use as much as possible free-of-charge, widely-used, secure, and open-source technologies. Deployment should take only a couple of days in a new environment, with minimal learning curve.
R6	The mobile data server can unify in-study data from the same participant (non-functional).	The mobile data server can pseudo-identify participants across their data sources (e.g., by assigning each participant a random identifier and separating it from personally identifiable data, such as a set of demographic data points). Data from multiple sources can be aligned in time. Internal identification is necessary to retrieve and delete information about participants.
R7	The mobile data server can support offline data (non-functional)	Mobile clients should be temporarily self-standing while collecting data in geographically remote experimental settings, and eventually synchronize with the server.
R8	The mobile data server can manage the participant data within safety requirements (non-functional).	Data management includes collection, storage, extraction, and analysis. These processes should be compliant with regulatory bodies, and data protection regulations. Security and privacy concerns are an important factor for participants to share their data [25, 18], e.g., by supporting participant consent, and their right to withdraw from the study at anytime.
	Requirements for the system	
R9	Researchers can manage separate features in studies with minimal programmatic changes (functional).	Researchers should be able to add, modify, and remove features for each study. Study editing should be possible while the study is in progress, to adapt to preliminary findings in the study. For example, researchers should be able to change survey questions, or swap IMWU devices. Researchers should be able to reuse features across multiple studies. Meeting this requirement enables quick iterations of hypothesis and deduction.
R10	Researchers can administer interventions (functional).	Researchers should be able to reach anonymous participants, either manually or automatically, potentially from real-time data analysis (e.g., by means of push notifications).
R11	Researchers can analyze in-study data (functional).	Researchers should be able to programmatically extract data for a study by using, e.g., queries. The data extracted should allow visualization, summarization, statistics, and machine learning processes, which can start on the platform and continue to a capacity limit which depends on the host environment. Researchers should be able to monitor participant retention and engagement with the study, to assess data quality.

Table 2: Overview of popular mobile sensing solutions and general characteristics

Name	Platform	Research Methods	Mobile Framework	Maintained[†]
AWARE, 2015 [16]	Android, iOS	Smartphone sensors, EMA	Yes	Yes
Sensus, 2013 [27]	Android, iOS	Smartphone sensors, EMA	Yes	Yes
SensingKit, 2016 [24]	Android, iOS	Smarphone sensors	No	Yes
Paco, 2014 [19]	Adnroid, iOS	EMA	No	Fair
Ohmage, 2015 [34]	Android, iOS	Smartphone sensors, EMA	No	No
Funf, 2011 [1]	Android	Smartphone sensors, EMA	No	No
Emotion Sense,2013 [27]	Android	Smartphone sensors, EMA	No	No
Research Stack, 2016 [8]	Android	EMA	No	No

[†] Self-assessed by the authors as of May 2020.

Existing platforms such as AWARE [16] suffer from high specificity in terms of their ability to integrate with other platforms, while others such as Sensus [27] are too stringent: customization involves considerable effort. Libraries such as SensingKit [24] cannot support mobile subject studies on their own. All are provided on an *as-is* basis, with

limited support and timely troubleshooting. Furthermore, libraries not actively maintained are at risk of obsolescence due to the rapid evolution of mobile operating systems.

This paper makes a unique contribution by sharing a set of instructions and guidance from our experience on numerous studies and experiments. We argue that researchers with long-term goals in mobile studies will benefit from creating a robust and flexible platform of their own, which can be modified in time to support changing research needs.

3. The *mQoL Lab* Platform

3.1. Architectural Design Overview

Conceptual Model: Considering the mobile data server of Fig. 1, each component (depicted in magenta), consists of two layers, robust and flexible, that we illustrate in Fig. 2. For each component, a set of permanent core features forms a robust layer (depicted in green). The features in the robust layer are connected across components (low coupling). The robust layer serves as the foundation of the architecture. In each component, a set of transient features (high cohesion) forms the flexible layer (depicted in yellow). By connecting and disconnecting the flexible features to their corresponding robust features in their component (high cohesion), researchers can adapt the platform to changing research needs. Fig. 2 illustrates this separation in layers, that all together fit the requirements in section 2.

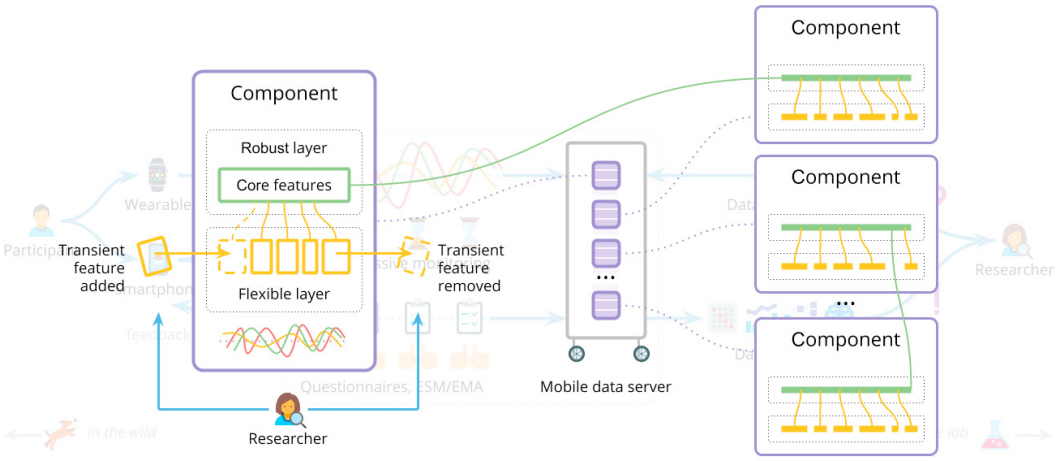


Fig. 2: Conceptual model of the *mQoL Lab* platform facilitating mobile subject studies

Architectural Design: The architectural design of the *mQoL Lab* platform focuses on the system constituent from Fig. 1, and is depicted in Fig. 3. This design consists of two blocks, clients and servers, grouping components (red rectangles for clients, and blue rectangles for servers), e.g., mobile apps and application server. Each component is tagged with the technology of *mQoL Lab* choice (e.g., `parse server`, `Android`). But researchers can leverage the architectural design even if they choose other technologies (see Section 3.2).

For example, the functionality of collecting TechRO data from passive sensing can be partitioned into a robust layer and a flexible layer (Fig. 2) as follows. Most wearable manufacturers (such as Fitbit and Withings implemented in *mQoL Lab*) require a reference to our platform in their Application Programmable Interface (APIs), which they use to (1) identify our platform in the informed consent, (2) authorize our platform for data collection, and (3) notify our platform about new data. Also, many wearable manufacturers use the same authorization protocol (OAuth 2.0), and communication style (REST [17]). The robust layer consists of a web client feature and a web server feature, collectively, a web application, for participants to choose the wearable type they own, and initiate the authorization. The web application (1) conducts authorization flows, (2) stores the grants used to collect the data, and (3) sends the data to the data storage component. The flexible layer includes device/manufacturer specific features. For example, servers implement pagination for data collected (a Fitbit API feature), or notification of new data (a Withings API feature). Currently, we are implementing data collection features for Garmin and Polar wearables as features in the flexible layer. These features reuse a minimally changed robust layer.

Supported Study Designs: The architecture in Fig 3 enables researchers to conduct short or longitudinal studies, inside or outside the lab, observational or including interventions, collecting passive or active data from participants,

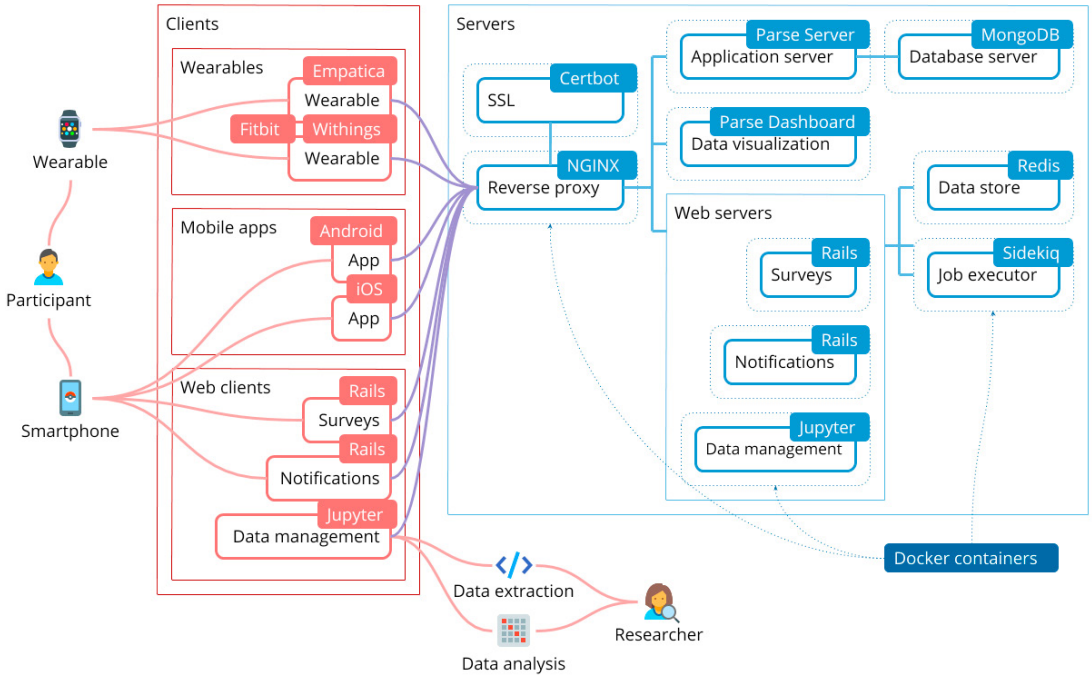


Fig. 3: Architecture of the *mQoL Lab* platform for mobile subject studies

including contextual markers. Table 3 describes study design and methodological aspects that researchers typically consider when selecting a platform. We describe how the the *mQoL Lab* architecture of Fig. 3 enables those aspects.

Table 3: Study design and methodological aspects supported by the *mQoL Lab* architectural

Design Aspect	**Description**
Study scope	The *mQoL Lab* design supports both observational (i.e., findings from data are used to further scientific knowledge) and intervention studies (i.e., findings can trigger interventions to one of more participants based on their collected data). Researchers can include Just-in-Time Adaptive Interventions (JITAIs) [32].
Study location	The *mQoL Lab* design allows both studies in the lab and in the wild. Participants can interact with IMWU devices in both cases.
Study variations	The *mQoL Lab* enables to create study flows with slight variations (e.g., with respect to data sources, frequency of data acquisition, among others). To this end for example, the *mQoL Lab* design allows the assignment of roles to study participants.
Study interactions	Human-device interactions in a study can range from standardized for all participants, to cross-sectional based on population characteristics, to personalized for a specific participant. The last two study interactions are based on data provided by participants.
Device provenance	Smartphones and wearable devices usuually follow standardized communication protocols. This allows the *mQoL Lab* design to support devices irrespective of their manufacturer.
Data collection source	Passive quantitative TechRO data can be collected by IMWU devices regularly carried by participants with them during daily life. The *mQoL Lab* for example includes a data logger for Android-based smartphones called *mQoL Log*. Active qualitative data is collected through mobile clients (such as web views in spartphone apps) where participants can answer surveys. Other connected devices, such as weight scales, can provide additional TechRO data.
Data synchronization	Some manufacturers provide data through an API (e.g. Fitbit, Withings),or via a mobile application through wireless transmission (e.g., NFC, BLE). The *mQoL Lab* design supports offline storage on mobile clients with eventual synchronization to the servers.
Data contextualization	Passive and active data can be augmented with contextual markers collected via *mQoL-log* such as time, location, ambient conditions, or social company (e.g., people around, collected either via wearables, smartphone sensors or self-reports).
Data sampling	Researchers can define sampling rates for passive and active data at the beginning or during a study. Sampling can be continuous or moment-, interval-, event-, and context-based, and it can be constrained to a given location area (geofenced).
Feedback to participants	Feedback can be given in near real-time as well as offline. It can be triggered locally (by using rules in the mobile clients) and remotely (automatically or manually, by the researcher).

3.2. Architectural Design and Implementation Choices

The architectural design of Fig. 3 uses a container-based infrastructure that simplifies the deployment, and allows execution as a distributed system, that supports the fulfillment of the requirements posed on the *mQoL Lab* platform. This choice provides several advantages: (1) the components can be maintained, updated, and run without influencing other components (R9) (cohesion), (2) they adhere to a common protocol of inter-container communication (coupling),

and (3) their deployment is reproducible onto any host which allows containerization (R5). For containerization, we selected Docker[1] for its ease of use, flexibility, ubiquity, and documentation. A docker image is a software package that includes all its dependencies (e.g. code, libraries, settings). In Fig. 3, each component is represented by a docker image, which describes a docker container able to communicate with other containers at run time. In the remainder of this section, we describe the fundamental components of the *mQoL Lab* platform.

The server block contains the application server component, a data visualization component, multiple web server components, and other components enabling data traffic and network security. The first server-side component is the *application server*, and it can be hosted locally (R4). It exposes the data in the platform as objects. The objects represent the entities used in the clients (e.g., survey questions and answers, passive sensor data), as well as those provided to the researchers via aggregation and data analysis tools (R6). It transfers them seamlessly between the clients and servers and stores them in the *database server*. The application server communicates with the clients and web servers. Furthermore, it allows for application logic hosted in the server to be triggered by the clients, simplifying the deployment of updates, and freeing up processing resources from the clients. For the application server we use Parse Server[2]. It represents data objects as (Parse Objects), includes software development kits (SDKs) to communicate with mobile applications, libraries to communicate with web servers (we use libraries for Rails/Ruby and Jupyter/Python) as well as the Cloud Code functionality to execute server logic (using Nodejs). For the database server, we chose MongoDB. The *data visualization* component has a dashboard with elevated permissions to manage the data as objects. We use Parse Dashboard which represents the data objects in the JSON format.

The server block contains the web server consisting of three web applications: (1) *surveys*, (2) *notifications*, and (3) *data management*. The operations of the web servers are optimized by an in-memory rapid *data store* (implemented using Redis) and a concurrent *job executor* (implemented using Sidekiq). The *surveys* web application renders dynamic surveys prepared for each mobile study. The surveys can be administered to the participant via conventional web clients (browsers), or through web views embedded in mobile applications (R1). The *notifications* web application schedules and dispatches push notifications to the appropriate clients and invites participants to answer the surveys (R10). Researchers can define the scheduling of each survey in a study by using a flexible scheme with time-based triggers [3]. We have chosen to implement the two web applications by using the Ruby on Rails web framework. The *data management* web application allows researchers to extract, summarize, aggregate, and visualize data by using a programmatic environment. They can analyze data by using Jupyter notebooks (R11) and libraries in R or Python. This last web application is currently an active area of research and development in the *mQoL Lab* platform.

The server-side network is managed through the *reverse proxy* and *SSL* components. The *reverse proxy* component routes network traffic from the internet directly into the various components. It is simple to use, integrates seamlessly with the other components, and allows the addition of new components (R9). For the reverse proxy component, we use nginx[4]. For secure communication over the web (R8), we use the SSL component. This component exposes a HTTPS certificate from within its container, which also executes letsencrypt[5] certificate authority.

In the clients block, our architectural design currently supports three types of clients: mobile applications, wearable devices, and web clients. Mobile applications use the Parse SDKs, which simplify the application logic and data transfer. The Parse SDKs allow local data storage on the clients prior to eventual synchronization (R7). For passive TechRO data collection, Android libraries allow efficient ways to send data from the on-board smartphone sensors to the application server directly (R4), in our case, via the *mQoL Log* component. For iOS, Apple Health Kit collects and stores the data locally on the device and exposes it to our mobile apps which send aggregates to the application server (R4). Mobile applications also embed web views of the surveys web client (R4). All mobile applications require informed consent from participants before starting the study (R2) and collecting data from any source (R3).

Wearables such as off-the-shelf fitness trackers (e.g., Fitbit, Withings) or research-oriented devices (e.g., Empatica) can be integrated into our platform. Two wearable data collection alternatives are supported at the moment. First, the manufacturer opens a web application where the participant can read and approve the informed consent for data usage

[1] https://www.docker.com/
[2] https://parseplatform.org/
[3] We provide an example in the public repository https://gitlab.unige.ch/qol/archimwu
[4] nginx is a web server, load balancer, and reverse proxy which requires minimum configuration: https://www.nginx.com/
[5] Certificate Authority provided by the Internet Security Research Group: https://letsencrypt.org/

(R3) and authorize our platform (via Oauth 2.0) to collect data from the wearable devices without revealing their credentials (R8). Second, the wearable device connects to the mobile application in the smartphone (via `Bluetooth`) and regularly synchronizes the data with it, in which case the consent in the mobile app covers the collected data (R3). Then the mobile applications eventually synchronize its data with the application server. The three web clients are part of the three web applications and communicate with their corresponding web servers.

We provide a public repository `https://gitlab.unige.ch/qol/archimwu` (*Repo*) with steps to bootstrap the foundation of a platform as shown in Fig. 3

3.3. Studies Conducted with the mQoL Lab Platform

Table 4 lists the research studies that leveraged various features of the *mQoL Lab* platform over time. The platform has been incrementally developed, enhanced, and it continues to be used in ongoing studies. The main features of the platform have remained since the early days, undergoing updates due to changes in the Android OS libraries. One novel functionality was added in the process to support an experimental method named Peer-ceived Momentary Assessment (PeerMA) which is being studied in the context of self and peer-based state assessments [2, 3].

Moreover, the process to re-instantiate the complete *mQoL Lab* platform has been validated at a new HIPAA compliant site in the USA during 2019. We accomplished this goal by following the steps outlined in *Repo*. Since then, one study was conducted to validate the stability of the platform, and the second, longitudinal study, used mobile technologies to assess quality of life-related aspects of patients undergoing a liver transplant.

Table 4: Studies conducted using the *mQoL Lab* platform

Study Aims	**Participants (N, t)**	**Methods and Tools**	**Year**	**Location**
Phone proximity [14]	28 x 1 month	DRM, EMA, Survey, Passive sensing	2011	USA
Mobile interaction experience [22]	29 x 1 month	DRM, EMA, Survey, Passive sensing	2012	USA
Intimacy perception [20]	20 x 1 month	DRM, EMA, Survey, Passive sensing	2013	Switzerland
Intimacy perception [21]	22 x 1 month	DRM, EMA, Survey, Passive sensing	2016	USA
Self-efficacy [36]	20 x 1 month	EMA, DRM, Survey, Passive sensing, Fitbit	2017	USA
Stress assessment [6]	25 x 1 month	EMA, Survey, Passive sensing	2018	Switzerland
Sleep assessment [7]	14 x 6 month	EMA, Survey, Passive sensing, Basis Peak	2018	Denmark
Sleep deprivation [10]	1 x 1 month	EMA, DRM, Survey, Fitbit, Glucose Monitor	2018	USA
Smartphone app quality of experience [12, 13]	38 x 1 month	EMA, Survey, Passive sensing	2018	Switzerland
Human state assessment with peers [2, 4]	30 x 1 month	EMA, PeerMA, Survey, Passive sensing	2018	USA, Switzerland
Health and dementia risk assessment [26]	20 x 3 months	Survey, Fitbit	2018	Denmark
Physical activity calibration [29]	31 x 2 years	Survey, Fitbit	2019	Denmark
Social support perception (active)	21 x 2 years	Survey, Fitbit	2019	Denmark
Quality of life in liver transplant patients (active)	15 x 6 months	EMA, PeerMA, Survey, Fitbit	2019	USA

4. Discussion and Concluding Remarks

In this paper, we characterized the main constituents of a mobile subjects study: participants, researchers, and the system. This is an area of active exploration and previous researchers have developed platforms, tools, and solutions to support it, especially those related to passively collecting data from wearables and smartphones. Given the fast pace of research, not all groups have the expertise or resources to design their own platform, and embracing inadequate frameworks, or siloed tools poses a high risk of obsolescence. Researchers with long-term goals in mobile sensing will benefit from building a reliable and scalable architecture that supports their growing needs. We described the *mQoL Lab* architecture that has evolved with more than ten studies over eight years; we focused not only on explaining the architecture, but also on the rationale of the underlying components of the architecture, offering practical and technical details that developers can use in the process of designing and building their platforms.

As future work, we plan to extend the contribution presented in this paper with videos, tutorials and code snippets that researchers can follow in a more hands-on manner with the aim of helping the community effectively.

Acknowledgements: We thank the following collaborators for their contributions: Mattia Gustarini, Frederik Bygballe, and Philip Lin. We also thank the Swiss FCS (2016-19), H2020 WellCo (no. 769765), GUARDIAN (2019-6-120-CP) and The University of Costa Rica.

References

[1] Aharony, N., Pan, W., Ip, C., Khayal, I., Pentland, A., 2011. Social fmri: Investigating and shaping social mechanisms in the real world. Pervasive Mob. Comput. 7, 643–659.

[2] Berrocal, A., Concepcion, W., De Dominicis, S., Wac, K., 2019. Complementing Human Behavior Assessment By Leveraging Personal Ubiquitous Devices and Social Links: Evaluation of the PeerMA Method. JMIR mHealth and uHealth (in press) .

[3] Berrocal, A., Wac, K., 2018. Peer-vasive computing: Leveraging peers to enhance the accuracy of self-reports in mobile human studies, in: MHC'18 Workshop. ACM Int. Conf. on Pervasive and Ubiquitous Computing, ACM, New York, NY, USA. pp. 600–605.

[4] Berrocal, A., Wac, K., 2020. Peers know you: A feasibility study of the predictive value of peer's observations to estimate human states, in: Int. Conf. on Mobile Systems and Pervasive Computing (to appear).

[5] Bongartz, H., Rübsamen, N., Raupach-Rosin, H., Akmatov, M., Mikolajczyk, R., 2017. Why do people participate in health-related studies? Int. journal of public health 62, 1059–1062.

[6] Ciman, M., Wac, K., 2016. Individuals' stress assessment using human-smartphone interaction analysis. IEEE T. on Affect. Comp. 9, 51–65.

[7] Ciman, M., Wac, K., 2018. mSleep: Individuals' Sleep Duration Assessment Using Human-Smartphone Interaction Analysis. Journal of Pervasive and Mobile Computing .

[8] Cornell, Accessed 01.15.2020. Researchstack: An sdk for building research study apps on android. http://researchstack.org/.

[9] Cornet, V., Holden, R., 2018. Systematic review of smartphone-based passive sensing for health and wellbeing.

[10] Daza, E., Wac, K., Oppezzo, M., 2019. Effects of Sleep Deprivation on Blood Glucose, Food Cravings, and Affect in a Non-Diabetic: An N-of-1 Randomized Pilot Study. Healthcare 8, 6.

[11] De Masi, A., Ciman, M., Gustarini, M., Wac, K., 2016. mQoL smart lab, in: ACM Int. Conf. on Pervasive and Ubiquitous Computing, ACM Press, New York, New York, USA. pp. 635–640.

[12] De Masi, A., Wac, K., 2018. You're Using This App For What? mQoL Living Lab Study, in: MHC Workshop at UbiComp'16.

[13] De Masi, A., Wac, K., 2019. Predicting quality of experience of popular mobile applications from a living lab study, in: Int. Conf. on Quality of Multimedia Experience, pp. 1–6.

[14] Dey, A., Wac, K., Ferreira, D., Tassini, K., Hong, J., Rojas, J., 2011. Getting Closer : An Empirical Investigation of the Proximity of User to Their Smart Phones, in: Int. Conf. on Ubiquitous Computing, Beijing, China.

[15] Dickert, N., Eyal, N., Goldkind, S., Grady, C., Joffe, S., Lo, B., Miller, F., Pentz, R., Silbergleit, R., Weinfurt, K., 2017. Reframing consent for clinical research: a function-based approach. The American Journal of Bioethics 17, 3–11.

[16] Ferreira, D., Kostakos, V., Dey, A., 2015. Aware: Mobile context instrumentation framework. Frontiers in ICT 2, 6.

[17] Fielding, R., 2000. Representational state transfer. Architectural Styles and the Design of Netowork-based Software Architecture , 76–85.

[18] Goodman, D., Bowen, D., Wenzel, L., Tehrani, P., Fernando, F., Khacheryan, A., Chowdhury, F., Johnson, C., Edwards, K., 2018. The research participant perspective related to the conduct of genomic cohort studies. Translational behavioral medicine 8, 119–129.

[19] Google, Accessed 01.15.2020. Paco: the personal analytics companion. https://www.pacoapp.com/.

[20] Gustarini, M., Wac, K., 2013. Smartphone Interactions Change for Different Intimacy Contexts, in: Int. Conf. on Mobile Computing, Applications and Services, Springer, Cham. pp. 72–89.

[21] Gustarini, M., Wac, K., Dey, A., 2016. Anonymous smartphone data collection: factors influencing the users' acceptance in mobile crowd sensing. Personal and Ubiquitous Computing 20.

[22] Ickin, S., Wac, K., Fiedler, M., Janowski, L., Hong, J., Dey, A., 2012. Factors influencing quality of experience of commonly used mobile applications. IEEE Communications Magazine , 48–56.

[23] Kahneman, D., Krueger, A., Schkade, D., Schwarz, N., Stone, A., 2004. The day reconstruction method. Science 306, 1776–1780.

[24] Katevas, K., Haddadi, H., Tokarchuk, L., 2016. Sensingkit: Evaluating the sensor power consumption in ios devices. IE'16 .

[25] Kerath, S., Klein, G., Kern, M., Shapira, I., Witthuhn, J., Norohna, N., Kline, M., Baksh, F., Gregersen, P., Taioli, E., 2013. Beliefs and attitudes towards participating in genetic research–a population based cross-sectional study. BMC Public Health 13, 114.

[26] Laghouila, S., Manea, V., Estrada, V., Wac, K., 2018. Digital health tools for chronic illness and dementia risk assessment in older adults, in: Annals of Behavioral Medicine, Oxford University Press. pp. S291–S291.

[27] Lathia, N., Rachuri, K., Mascolo, C., Roussos, G., 2013. Open source smartphone libraries for computational social science. UbiComp'13 .

[28] Lawson, A., 2015. Hypothetico-deductive Method. Springer Netherlands, Dordrecht. pp. 471–472.

[29] Manea, V., Berrocal, A., Wac, K., 2020. Using consumer-friendly wearables to associate patient- and technology-reported physical activity in healthy seniors, in: Int. Conf. on Mobile Systems and Pervasive Computing (to appear).

[30] Mayo, N., Figueiredo, S., Ahmed, S., Bartlett, S., 2017. Montreal Accord on patient-reported outcomes use series—paper 2: terminology proposed to measure what matters in health. Journal of Clinical Epidemiology 89, 119–124.

[31] Moller, A., Merchant, G., Conroy, D., West, R., Hekler, E., Kugler, K., Michie, S., 2017. Applying and advancing behavior change theories and techniques in the context of a digital health revolution. Journal of behavioral medicine 40, 85–98.

[32] Nahum-shani, I., Smith, S., Witkiewitz, K., Collins, L., Spring, B., Murphy, S., 2014. Just-in-time adaptive interventions (JITAIs): An organizing framework for ongoing health behavior support. The Methodology Center Technical Report 073975, 1–37.

[33] Stone, A., Shiffman, S., 1994. Ecological Momentary Assessment in Behavioral Medicine. Annals of Behavioral Medicine 16, 199–202.

[34] Tangmunarunkit, H., Hsieh, C., Longstaff, B., Nolen, S., Jenkins, J., Ketcham, C., Selsky, J., Alquaddoomi, F., George, D., et al., 2015. Ohmage: A general and extensible end-to-end participatory sensing platform. ACM Trans. Intelligent Syst. and Tech. 6, 1–21.

[35] Wac, K., 2018. From Quantified Self to Quality of Life. Springer International Publishing, Cham. pp. 83–108.

[36] Wulfovich, S., Fiordelli, M., Rivas, H., Concepcion, W., Wac, K., 2019. "I Must Try Harder": Design Implications for Mobile Apps and Wearables Contributing to Self-Efficacy of Patients With Chronic Conditions. Frontiers in Psychology 10, 1–12.

Publication 5: mQoL: Mobile Quality of Life Lab

Vlad Manea, Vero Estrada-Galiñanes, Katarzyna Wac. *mQoL: Mobile Quality of Life lab*. Poster and demo at the Digital Health Conference (DH 2018). **Nominated for the Innovation Prize in the category of the best data-driven innovation, Lyon, France**.

Abstract Chronic diseases are the top contributor to mortality worldwide. Their risk decreases with a healthy lifestyle, determined by daily life behavior. Reference chronic illness risk models combine patient performance-based reports with self-reports. Performance reports, obtained at the doctor's office, are collected with clinically approved devices to ensure high accuracy, but the process is expensive, momentary, and occurs outside of the patient's daily life. Self-reports are affordable, can be contextual and recur in time, but introduce perception bias and are prone to socially acceptable answers. Meanwhile, the market for mHealth, i.e., personalized devices that monitor daily life through behavioral markers (e.g., exercise or sleep), is gaining acceptance. Although many do not promise medical accuracy yet, the sheer data collected may provide useful insights into patterns. A caveat is that the mHealth space is fragmented. Numerous researchers design, build, deploy, and maintain applications that focus on a single experiment, a marker, or a disease. Such apps prove too narrow to address participant Quality of Life from a holistic perspective. In addition, researchers report back gaps in use (and data) from the apps. Faced with these data collection challenges ourselves and aiming at holistic QoL assessment, we are operationalizing the Mobile Quality of Life Lab. Our app aims at serving as an ecosystem of digital health exploration for participants and researchers. With personalized, contextual, and graphical content for participants to monitor, observe, and reflect upon daily life as a whole, we hypothesize that it can serve as a useful tool to make sense of behavior and life quality and potentially enable behavior change in the long term. By reusing its data collection and exploration apparatus on top of well-known frameworks such as ResearchKit, HealthKit, Open mHealth, or AWARE, our app enables to obtain performance-reported outcomes by measurement of behavioral markers in time and context. Also, we include general self-reports on demographics and Quality of Life as well as domain-specific self-reports in the explorations. Researchers like ourselves can now focus only on collecting approved, consented, anonymized, contextual, and chronologic data, while providing participants timely, personalized, contextual, and beneficial information from their investigations. We intend to leverage the app in Europe within the WellCo H2020 project where we aim to manage cardiovascular disease risk. We have 300 participants planned for 2018.

Keywords mobile application, mobile platform, mixed methods, passive sensing, data collection, behaviours, quality of life.

QoLife Technologies

Vlad Manea, Veronica Estrada, Katarzyna Wac
manea@di.ku.dk • www.qol.unige.ch

European Network of Living Labs

mQoL: Mobile Quality of Life Lab

Mobile application for Quality of Life assessment and longitudinal behavioral data collection in time and context.

Researchers

Run custom explorations

Collect behavioral data

Avoid building a new app

Participants

Monitor daily life

Change behaviors?

Reduce disease risk?

Data

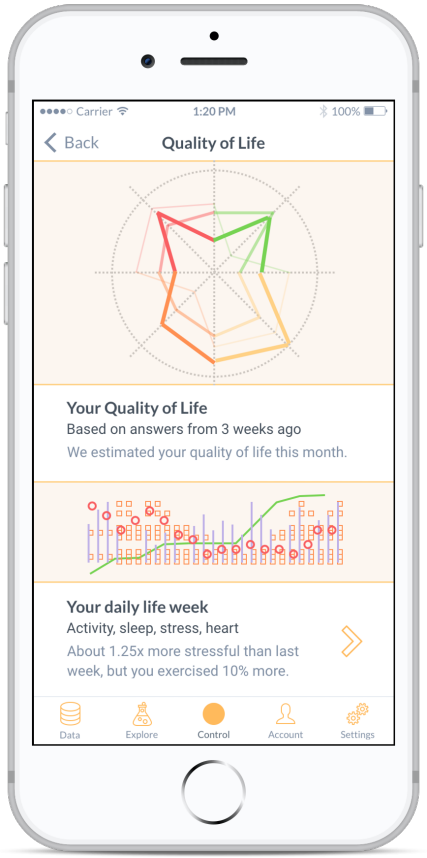
Performance-reported

- *behavioral markers*
- *device use*

Self-reported

- *Quality of Life*
- *demographic*
- *exploration-specific*

Let's talk about your next experiments!



UNIVERSITÉ DE GENÈVE

UNIVERSITY OF COPENHAGEN

EUROPEAN SCIENCE FOUNDATION

cost EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

EUROPEAN COMMISSION

HORIZON 2020

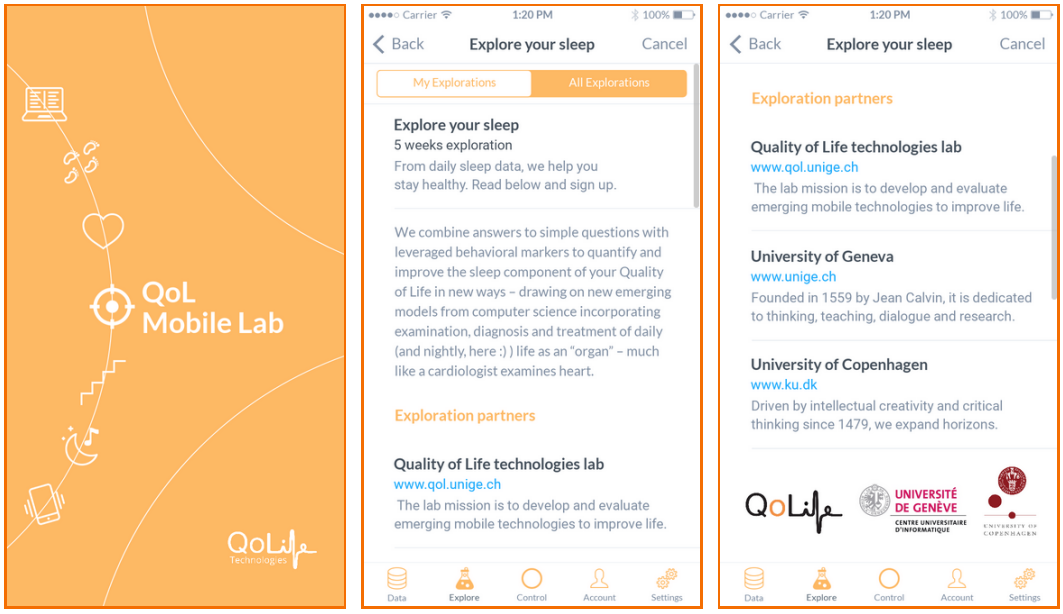
AAL AMBIENT ASSISTED LIVING

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

FNSNF

FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

Research at Google



(a) Mobile Quality of Life Lab (mQoL).

(b) Sign up for a sleep exploration. Part I: background.

(c) Sign up for a sleep exploration. Part II: researchers.

(d) Sign up for a sleep exploration. Part III: existing scientific evidence.

(e) Sign up for a sleep exploration. Part IV: instructions to provide PROs.

(f) Sign up for a sleep exploration. Part V: instructions to collect TechROs.

Fig. A.1.: Excerpts from the mQoL mobile app design. Part of Publication 5 which was demonstrated. The app design was not peer-reviewed.

(g) Overview of insights from all ongoing explorations.

(h) Overview of insights from the sleep exploration.

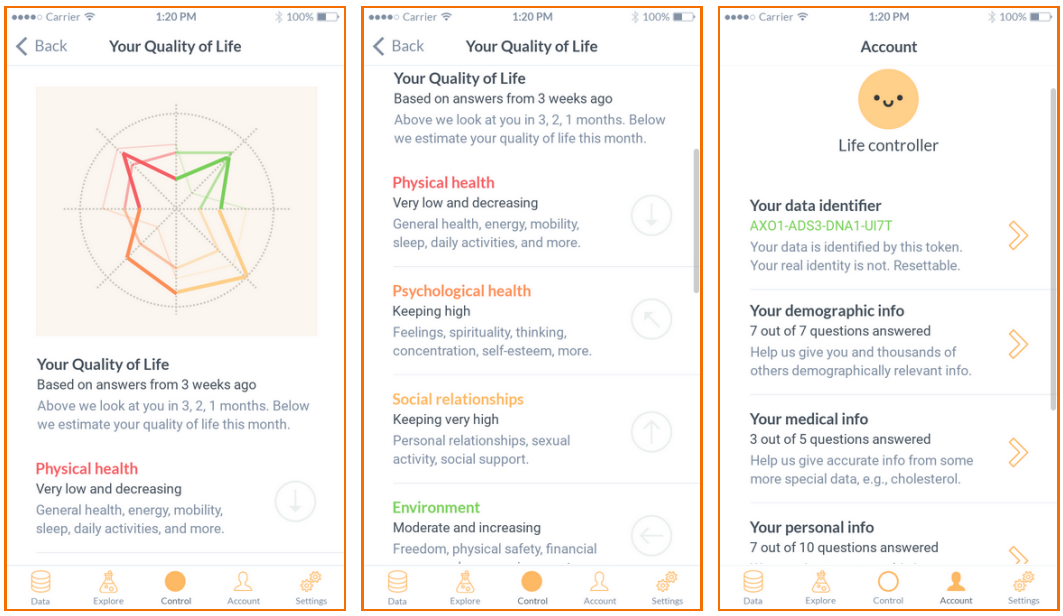
(i) PRO data collection for the sleep exploration.

(j) TechRO data collection for the sleep exploration.

(k) Data control options for the sleep exploration.

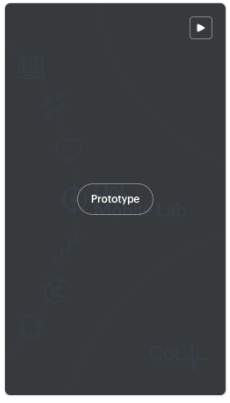
(l) Control of participation with notification *cards*.

Fig. A.1.: Excerpts from the mQoL mobile app design (continued). Part of Publication 5 which was demonstrated. The app design was not peer-reviewed.



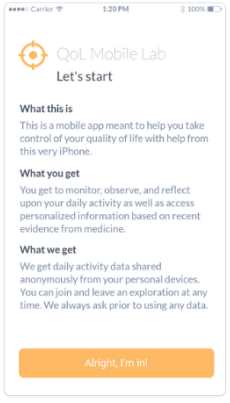
(m) Quality of Life view. Part I. **(n)** Quality of Life view. Part II. **(o)** Pseudonymous user account.

Fig. A.1.: Excerpts from the mQoL mobile app design (continued). Part of Publication 5 which was demonstrated. The app design was not peer-reviewed. The complete app design continues on the following pages.



Updated 22nd Aug @ 00:31 AM

Intro



Updated 2nd Mar @ 10:41 AM

Welcome

Updated 28th Apr @ 21:58 PM

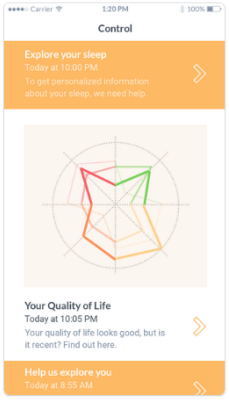
Consent

Updated 28th Apr @ 21:55 PM

Consent : Agreed age

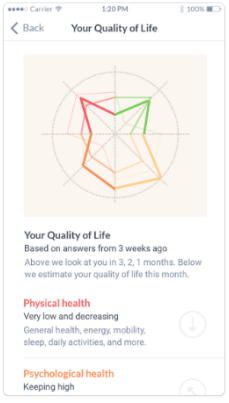
Updated 28th Apr @ 21:55 PM

Consent : Agreed willfulness



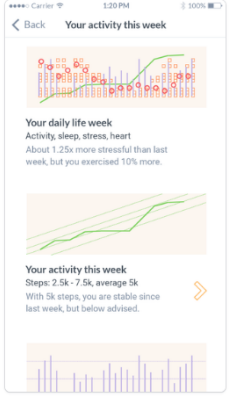
Updated 22nd Aug @ 00:38 AM

Control



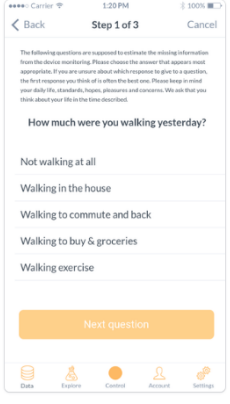
Updated 22nd Aug @ 00:51 AM

Control > Your Quality of Life



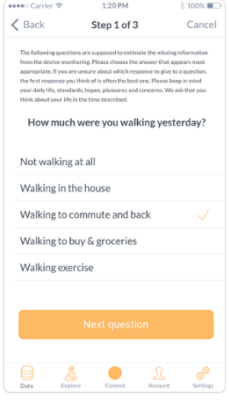
Updated 22nd Aug @ 00:52 AM

Control > Your Activity this week



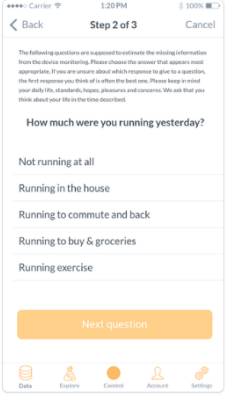
Updated 2nd Mar @ 10:56 AM

Control > Help us fill in the gaps > Step 1



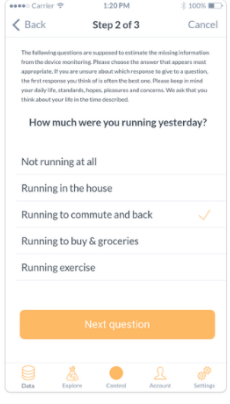
Updated 8th Feb @ 16:21 PM

Control > Help us fill in the gaps > Step 1 : Walking to commute and back



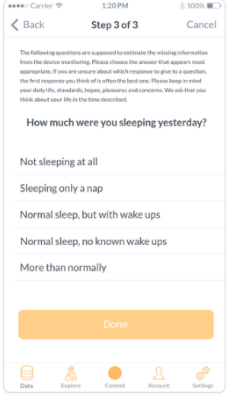
Updated 2nd Mar @ 10:57 AM

Control > Help us fill in the gaps > Step 2



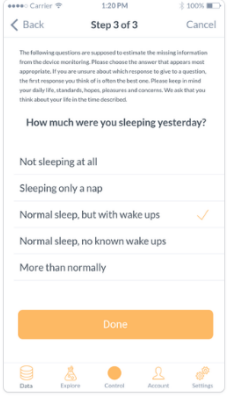
Updated 8th Feb @ 16:22 PM

Control > Help us fill in the gaps > Step 2 : Running to commute and back



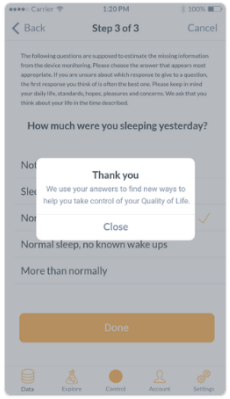
Updated 2nd Mar @ 10:58 AM

Control > Help us fill in the gaps > Step 3



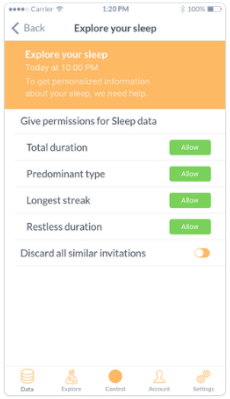
Updated 8th Feb @ 16:22 PM

Control > Help us fill in the gaps > Step 3 : Normal sleep, but with wake ups



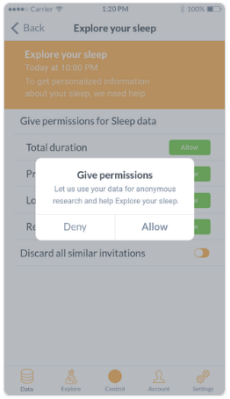
Updated 2nd Mar @ 10:59 AM

Control > Help us fill in the gaps > Step 3 : Thank you



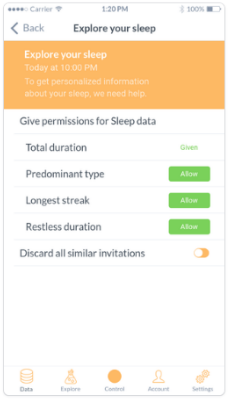
Updated 2nd Mar @ 11:01 AM

Control > Explore your sleep



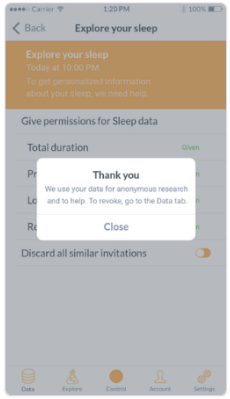
Updated 2nd Mar @ 11:01 AM

Control > Explore your sleep : Give permissions



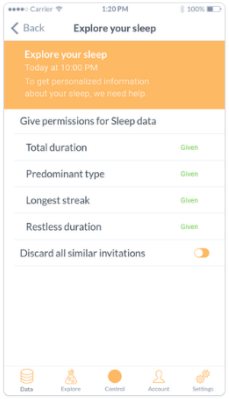
Updated 22nd Feb @ 00:20 AM

Control > Explore your sleep : Given permissions



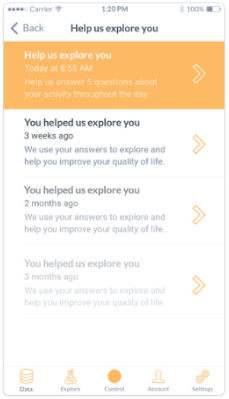
Updated 2nd Mar @ 11:03 AM

Control > Explore your sleep : Giving all permissions thank you message



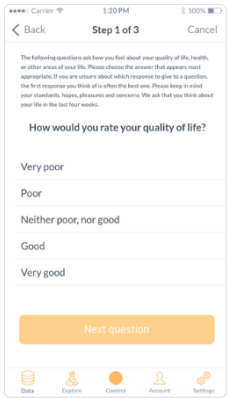
Updated 22nd Feb @ 00:20 AM

Control > Explore your sleep : Given all permissions



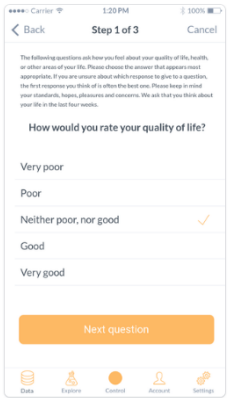
Updated 18th Mar @ 21:15 PM

Control > Help us explore you



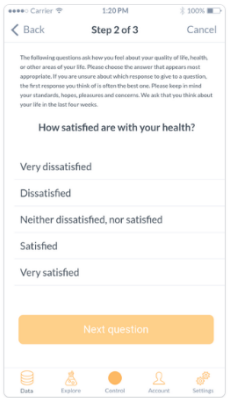
Updated 2nd Mar @ 11:09 AM

Control > Help us explore you > Step 1



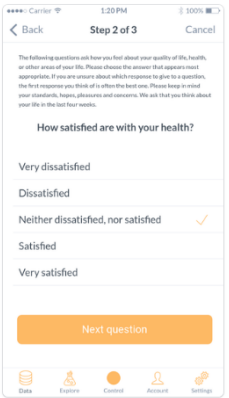
Updated 8th Feb @ 13:07 PM

Control > Help us explore you > Step 1 : Neither poor, nor good



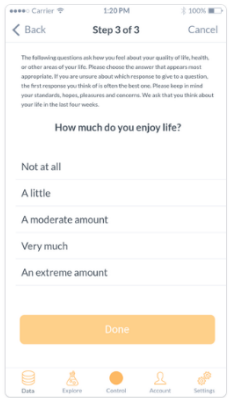
Updated 2nd Mar @ 11:10 AM

Control > Help us explore you > Step 2



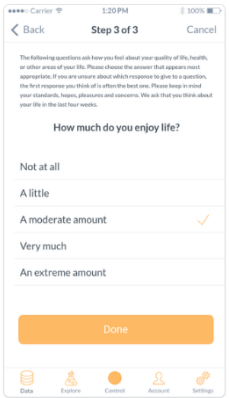
Updated 8th Feb @ 12:55 PM

Control > Help us explore you > Step 2 : Neither dissatisfied, nor satisfied



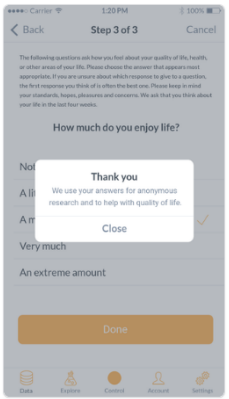
Updated 2nd Mar @ 11:10 AM

Control > Help us explore you > Step 3



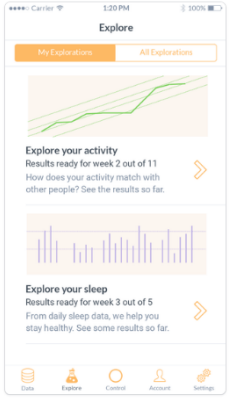
Updated 8th Feb @ 12:54 PM

Control > Help us explore you > Step 3 : A moderate amount



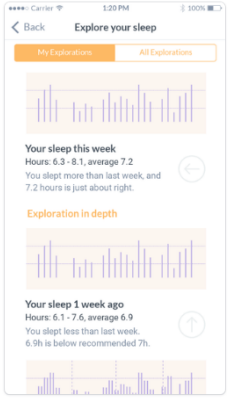
Updated 2nd Mar @ 11:23 AM

Control > Help us explore you > Step 3 : Thank you



Updated 22nd Jul @ 11:50 AM

Explore > My Explorations

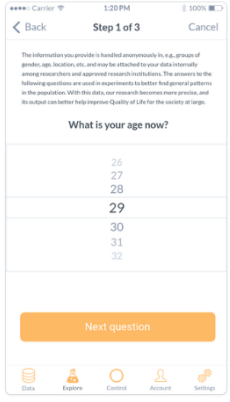


Updated 22nd Jul @ 11:50 AM

Explore > My Explorations > Explore your sleep

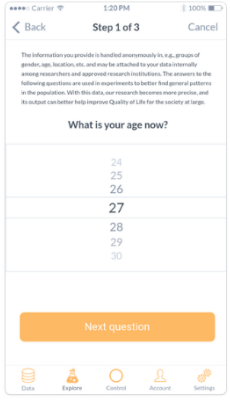
Updated 2nd Mar @ 11:32 AM

Explore > My Explorations > Explore your sleep > Step 1



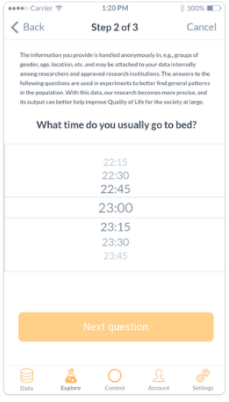
Updated 11th Feb @ 10:23 AM

Explore > My Explorations > Explore your sleep > Step 1 : Swiped age up



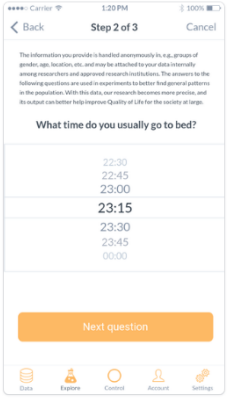
Updated 11th Feb @ 10:23 AM

Explore > My Explorations > Explore your sleep > Step 1 : Swiped age down



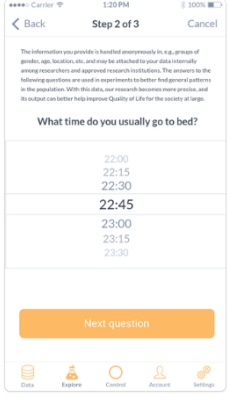
Updated 2nd Mar @ 11:33 AM

Explore > My Explorations > Explore your sleep > Step 2



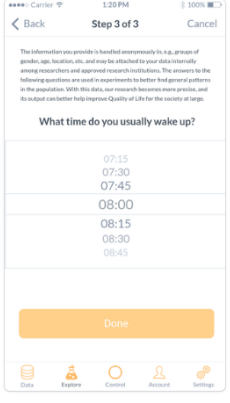
Updated 11th Feb @ 10:26 AM

Explore > My Explorations > Explore your sleep > Step 2 : Swiped wake up time up



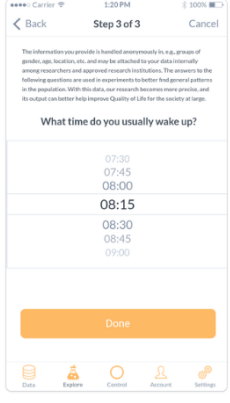
Updated 11th Feb @ 10:25 AM

Explore > My Explorations > Explore your sleep > Step 2 : Swiped wake up time down



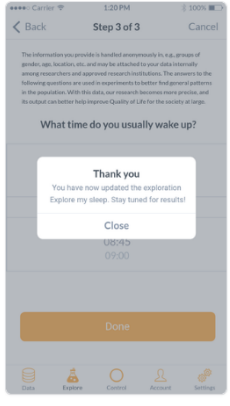
Updated 2nd Mar @ 11:39 AM

Explore > My Explorations > Explore your sleep > Step 3



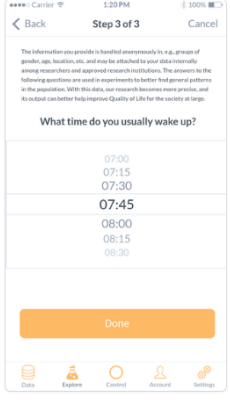
Updated 11th Feb @ 10:27 AM

Explore > My Explorations > Explore your sleep > Step 3 : Swiped wake up time up



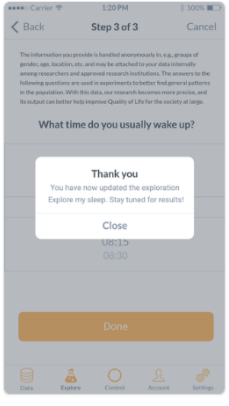
Updated 2nd Mar @ 11:39 AM

Explore > My Explorations > Explore your sleep > Step 3 : Swiped wake up time up thank you message



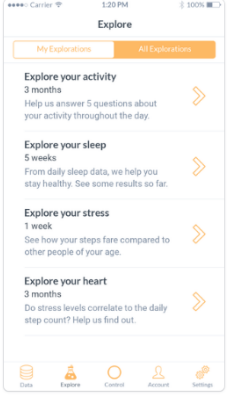
Updated 11th Feb @ 10:36 AM

Explore > My Explorations > Explore your sleep > Step 3 : Swiped wake up time down



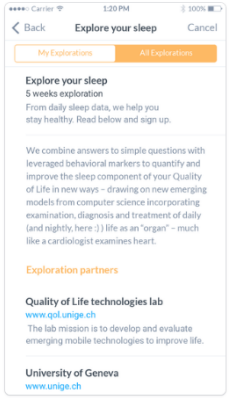
Updated 19th Feb @ 21:56 PM

Explore > My Explorations > Explore your sleep > Step 3 : Swiped wake up time down thank you message



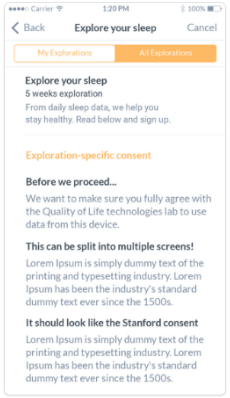
Updated 3rd Jun @ 12:40 PM

Explore > All Explorations



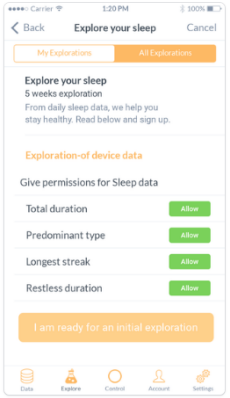
Updated 22nd Jul @ 11:51 AM

Explore > All Explorations > Explore your sleep : Overview before signup



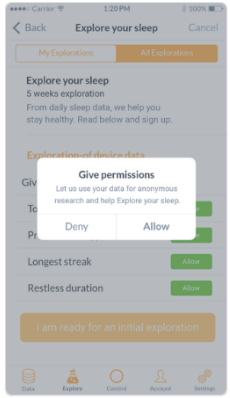
Updated 22nd Jul @ 11:52 AM

Explore > All Explorations > Explore your sleep : Consent



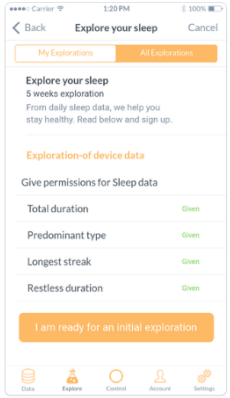
Updated 22nd Jul @ 11:52 AM

Explore > All Explorations > Explore your sleep : Permissions



Updated 2nd Mar @ 11:49 AM

Explore > All Explorations > Explore your sleep : Giving permissions



Updated 10th Feb @ 20:50 PM

Explore > All Explorations > Explore your sleep : Given permissions

Updated 2nd Mar @ 11:52 AM

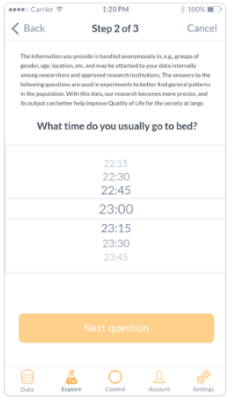
Explore > All Explorations > Explore your sleep > Step 1

Updated 10th Feb @ 21:07 PM

Explore > All Explorations > Explore your sleep > Step 1 : Swiped age up

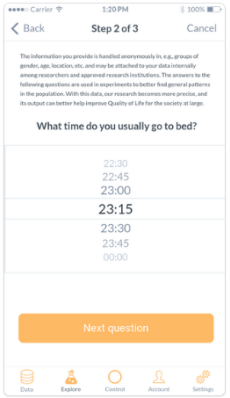
Updated 10th Feb @ 21:37 PM

Explore > All Explorations > Explore your sleep > Step 1 : Swiped age down



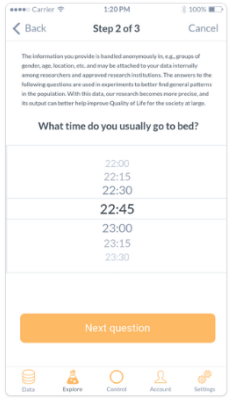
Updated 2nd Mar @ 11:52 AM

Explore > All Explorations > Explore your sleep > Step 2



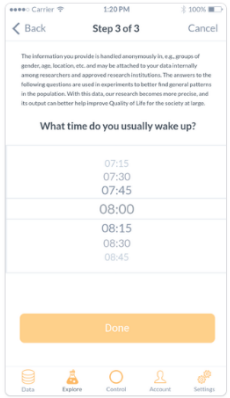
Updated 10th Feb @ 21:26 PM

Explore > All Explorations > Explore your sleep > Step 2 : Swiped wake up time up



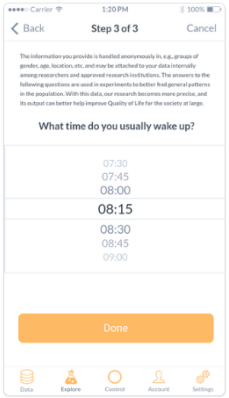
Updated 10th Feb @ 21:26 PM

Explore > All Explorations > Explore your sleep > Step 2 : Swiped wake up time down



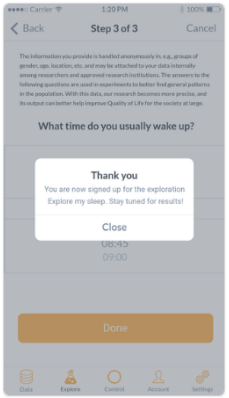
Updated 2nd Mar @ 11:53 AM

Explore > All Explorations > Explore your sleep > Step 3



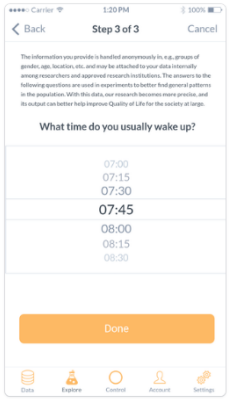
Updated 10th Feb @ 21:38 PM

Explore > All Explorations > Explore your sleep > Step 3 : Swiped wake up time up



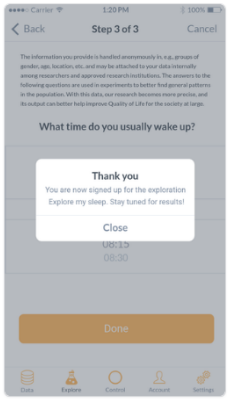
Updated 2nd Mar @ 11:53 AM

Explore > All Explorations > Explore your sleep > Step 3 : Swiped wake up time up thank you message



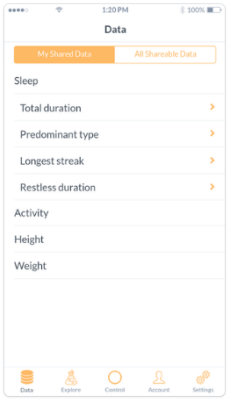
Updated 10th Feb @ 21:39 PM

Explore > All Explorations > Explore your sleep > Step 3 : Swiped wake up time down



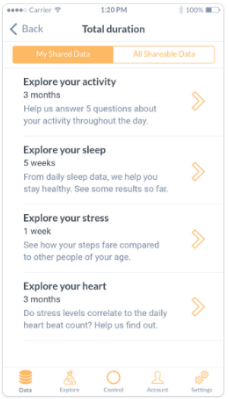
Updated 19th Feb @ 21:57 PM

Explore > All Explorations > Explore your sleep > Step 3 : Swiped wake up time down thank you message



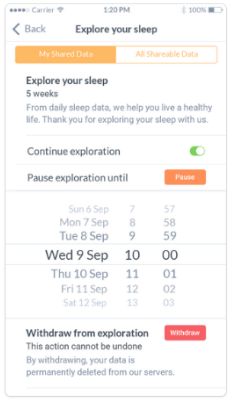
Updated 2nd Mar @ 11:55 AM

Data > My Shared Data



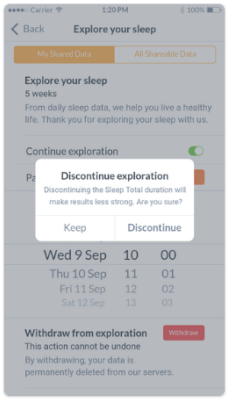
Updated 2nd Mar @ 11:56 AM

Data > My Shared Data > Sleep Total Duration



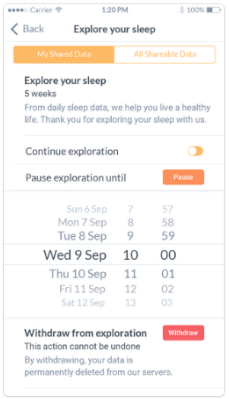
Updated 22nd Aug @ 00:56 AM

Data > My Shared Data > Sleep Total Duration > Explore your sleep



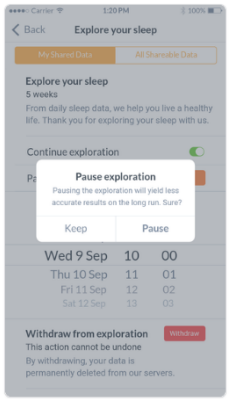
Updated 2nd Mar @ 12:18 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Discontinue exploration



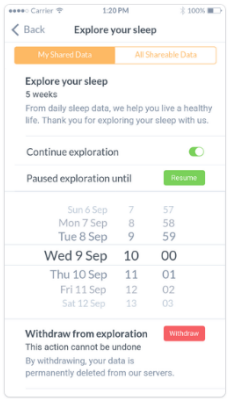
Updated 17th Apr @ 21:03 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Discontinued exploration



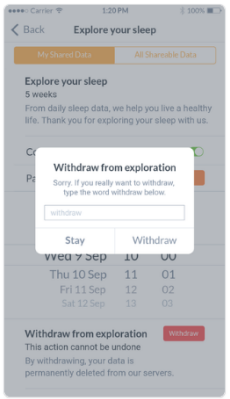
Updated 2nd Mar @ 12:19 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Pausing exploration



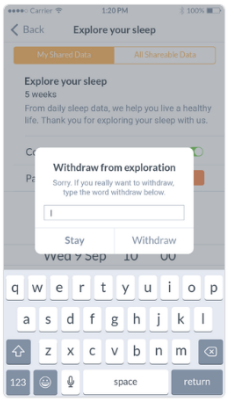
Updated 10th Feb @ 22:12 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Paused exploration



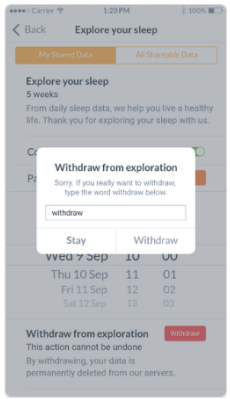
Updated 2nd Mar @ 12:19 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Withdrawing from exploration



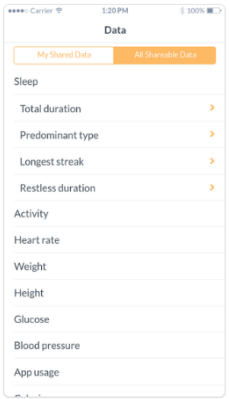
Updated 11th Feb @ 13:10 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Withdrawing from exploration confirming



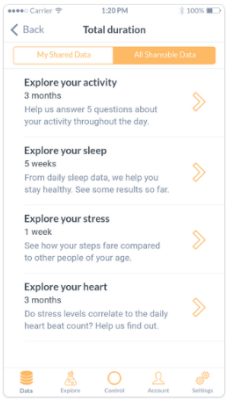
Updated 10th Feb @ 22:19 PM

Data > My Shared Data > Sleep Total Duration > Explore your sleep : Withdraw from exploration last step



Updated 2nd Mar @ 12:24 PM

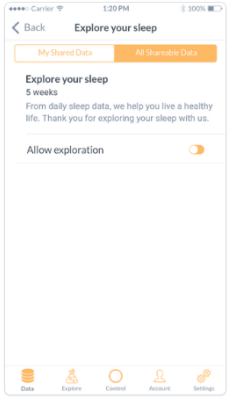
Data > All Shareable Data



Updated 2nd Mar @ 12:24 PM

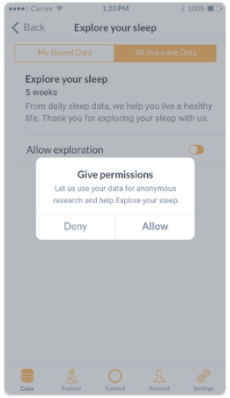
Data > All Shared Data > Sleep Total Duration

confirming



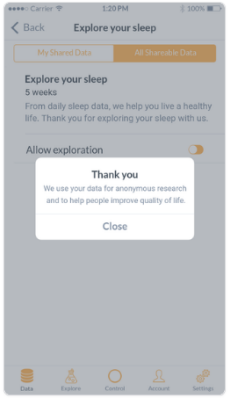
Updated 2nd Mar @ 12:24 PM

Data > All Shared Data > Sleep Total Duration > Explore your sleep



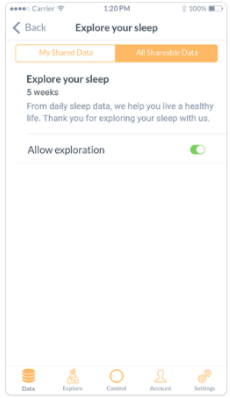
Updated 2nd Mar @ 12:25 PM

Data > All Shared Data > Sleep Total Duration > Explore your sleep : Give permissions



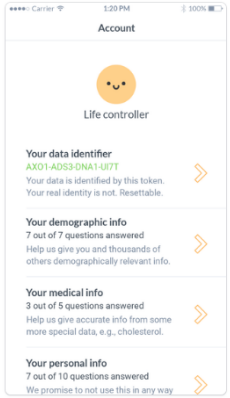
Updated 2nd Mar @ 12:25 PM

Data > All Shared Data > Sleep Total Duration > Explore your sleep : Given permissions thank you message



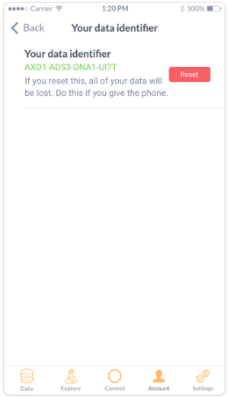
Updated 9th Feb @ 10:55 AM

Data > All Shared Data > Sleep Total Duration > Explore your sleep : Given permissions



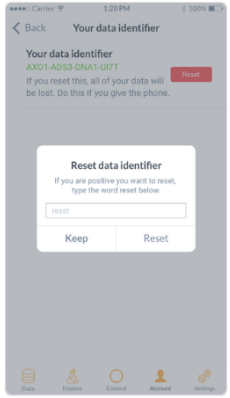
Updated 22nd Aug @ 00:56 AM

Account



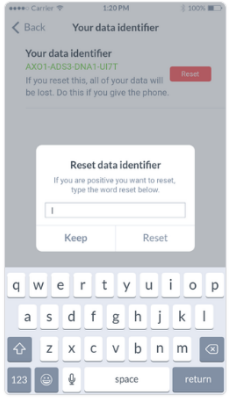
Updated 2nd Mar @ 12:37 PM

Account > Your data identifier



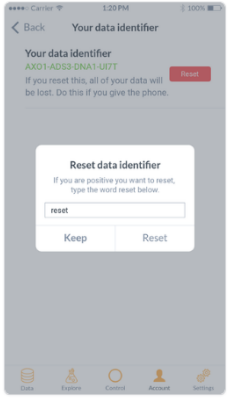
Updated 2nd Mar @ 12:37 PM

Account > Your data identifier : Reset warning prompt



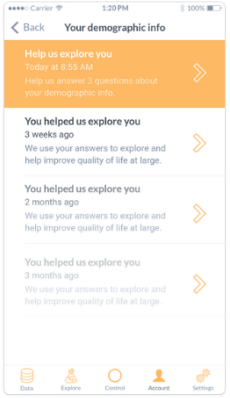
Updated 10th Feb @ 22:40 PM

Account > Your data identifier : Reset warning typing reset



Updated 10th Feb @ 22:40 PM

Account > Your data identifier > Reset warning confirming



Updated 2nd Mar @ 18:11 PM

Account > Your demographic info

Updated 2nd Mar @ 18:12 PM

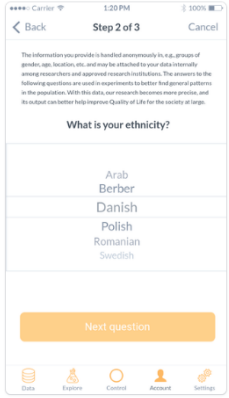
Account > Your demographic info > Step 1

Updated 2nd Mar @ 18:12 PM

Account > Your demographic info > Step 1 : Swiped age down

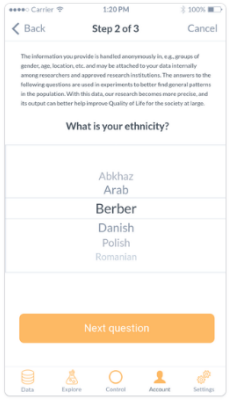
Updated 2nd Mar @ 18:12 PM

Account > Your demographic info > Step 1 : Swiped age up



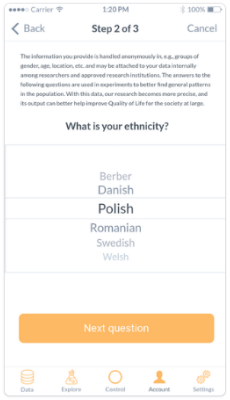
Updated 2nd Mar @ 18:13 PM

Account > Your demographic info > Step 2



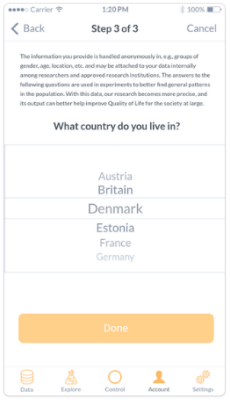
Updated 2nd Mar @ 18:13 PM

Account > Your demographic info > Step 2 : Swiped ethnicity down



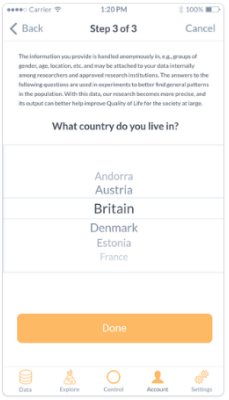
Updated 2nd Mar @ 18:18 PM

Account > Your demographic info > Step 2 : Swiped ethnicity up



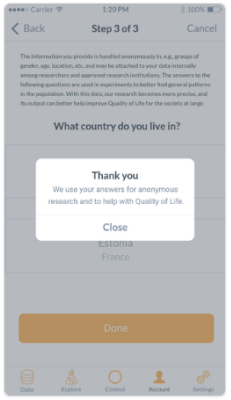
Updated 2nd Mar @ 18:13 PM

Account > Your demographic info > Step 3



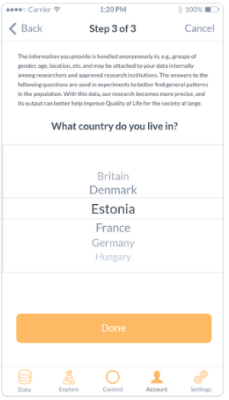
Updated 2nd Mar @ 18:13 PM

Account > Your demographic info > Step 3 : Swiped country down



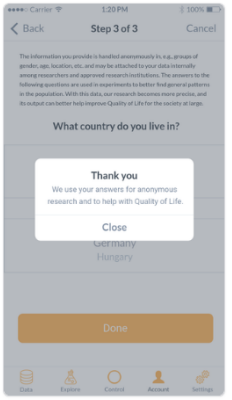
Updated 2nd Mar @ 18:21 PM

Account > Your demographic info > Step 3 : Swiped country down : Thank you message



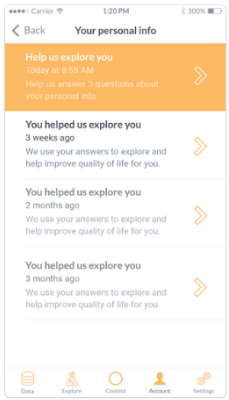
Updated 2nd Mar @ 18:14 PM

Account > Your demographic info > Step 3 : Swiped country up



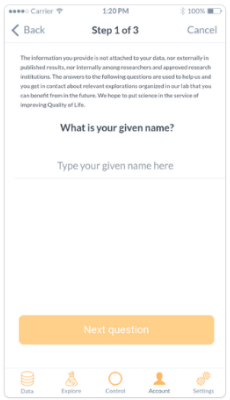
Updated 2nd Mar @ 18:21 PM

Account > Your demographic info > Step 3 : Swiped country up : Thank you message



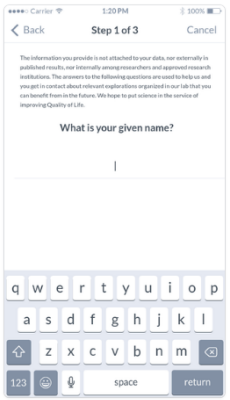
Updated 2nd Mar @ 18:15 PM

Account > Your personal info



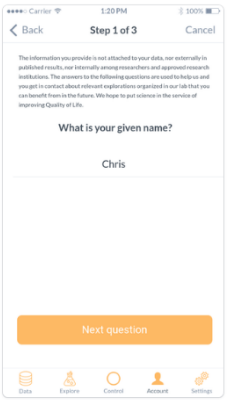
Updated 2nd Mar @ 18:32 PM

Account > Your personal info > Step 1



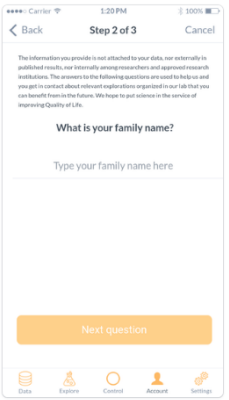
Updated 2nd Mar @ 18:32 PM

Account > Your personal info > Step 1 : Typing first name



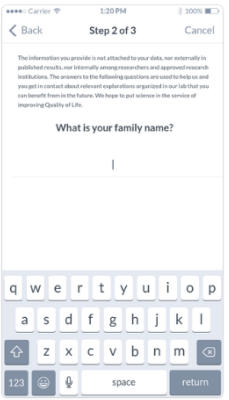
Updated 2nd Mar @ 18:33 PM

Account > Your personal info > Step 1 : Typed first name



Updated 2nd Mar @ 18:33 PM

Account > Your personal info > Step 2

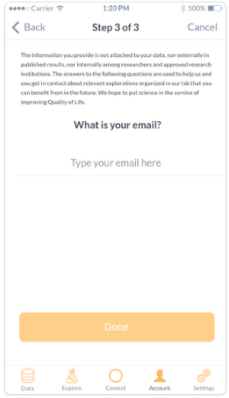


Updated 2nd Mar @ 18:33 PM

Account > Your personal info > Step 1 : Typing last name

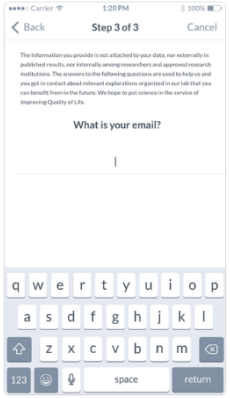
Updated 2nd Mar @ 18:34 PM

Account > Your personal info > Step 1 : Typed last name



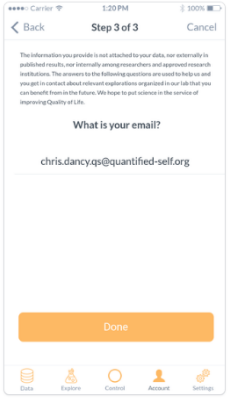
Updated 2nd Mar @ 18:34 PM

Account > Your personal info > Step 3



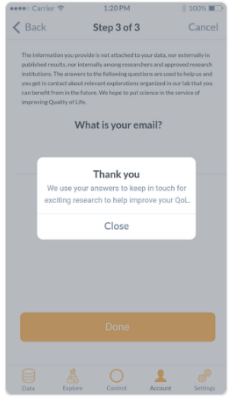
Updated 2nd Mar @ 18:35 PM

Account > Your personal info > Step 3 : Typing email



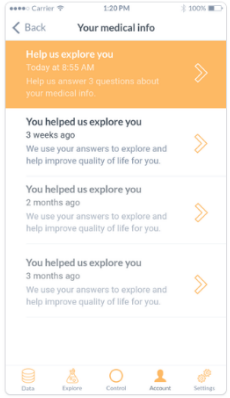
Updated 2nd Mar @ 18:35 PM

Account > Your personal info > Step 3 : Typed email



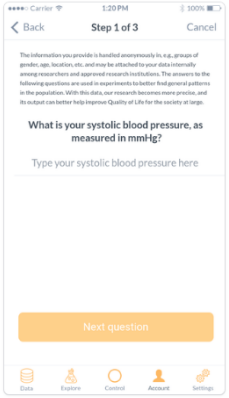
Updated 2nd Mar @ 18:35 PM

Account > Your personal info > Step 3 : Thank you message



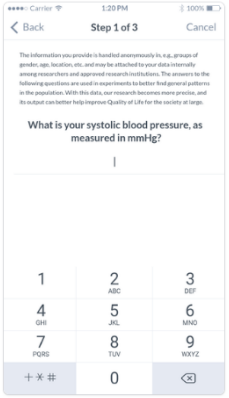
Updated 2nd Mar @ 19:23 PM

Account > Your medical info



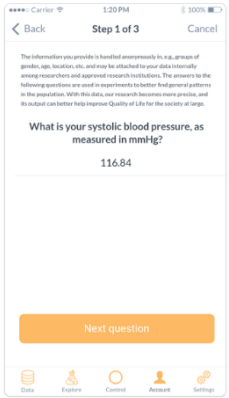
Updated 2nd Mar @ 19:21 PM

Account > Your medical info > Step 1



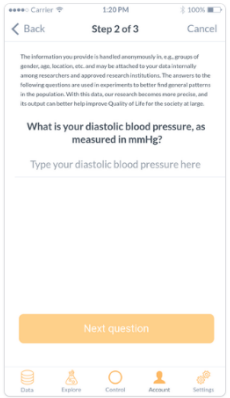
Updated 2nd Mar @ 19:08 PM

Account > Your medical info > Step 1 : Typing systolic blood pressure



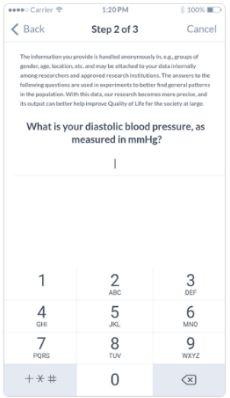
Updated 2nd Mar @ 19:08 PM

Account > Your medical info > Step 1 : Typed systolic blood pressure



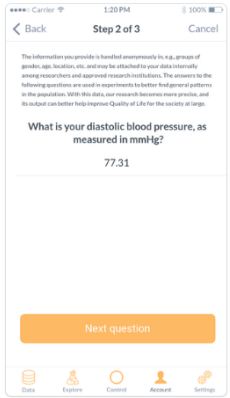
Updated 2nd Mar @ 19:21 PM

Account > Your medical info > Step 2



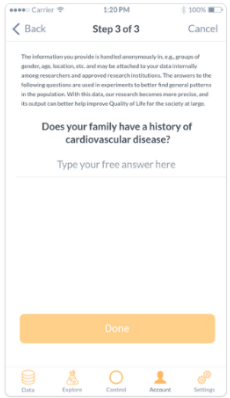
Updated 2nd Mar @ 18:56 PM

Account > Your medical info > Step 1 : Typing diastolic blood pressure



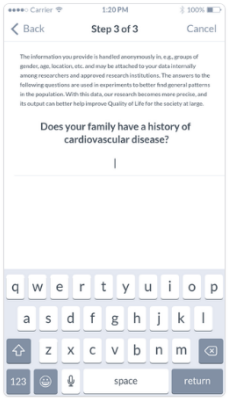
Updated 2nd Mar @ 18:56 PM

Account > Your medical info > Step 1 : Typed diastolic blood pressure



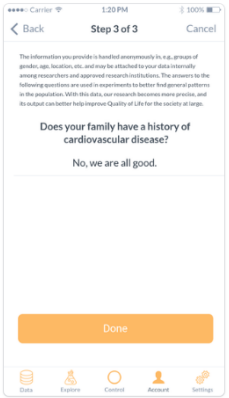
Updated 2nd Mar @ 19:21 PM

Account > Your medical info > Step 3



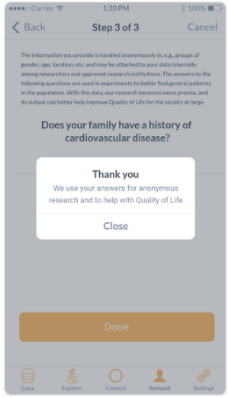
Updated 2nd Mar @ 18:55 PM

Account > Your medical info > Step 3 : Typing family risk



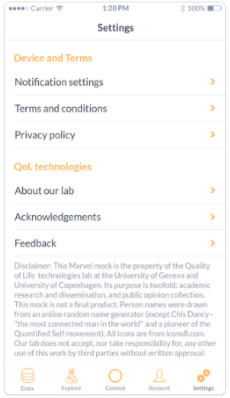
Updated 2nd Mar @ 18:58 PM

Account > Your medical info > Step 3 : Typed family risk



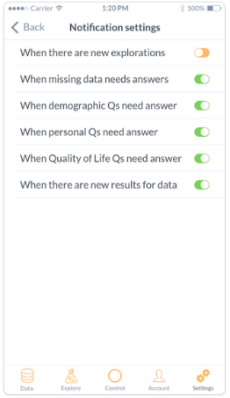
Updated 2nd Mar @ 19:22 PM

Account > Your medical info > Step 3 : Thank you message



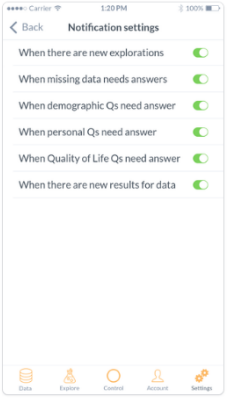
Updated 18th Mar @ 22:04 PM

Settings



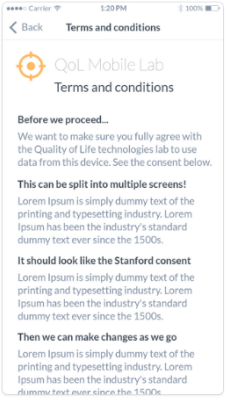
Updated 2nd Mar @ 12:52 PM

Settings > Notification settings



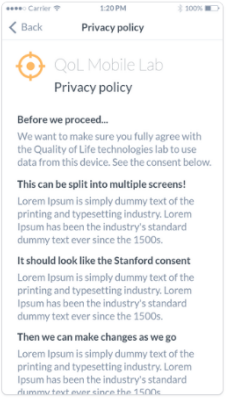
Updated 10th Feb @ 15:51 PM

Settings > Notification settings : New explorations enabled



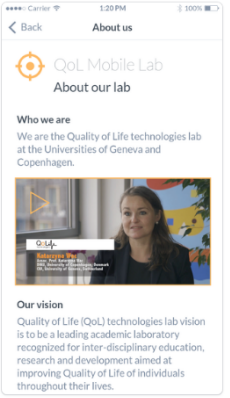
Updated 2nd Mar @ 12:52 PM

Settings > Terms and conditions



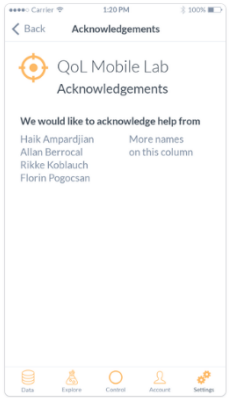
Updated 2nd Mar @ 12:53 PM

Settings > Privacy policy



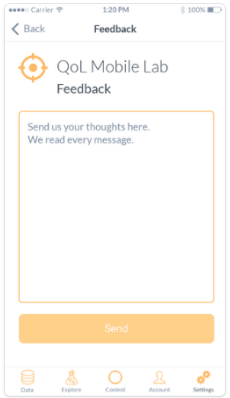
Updated 2nd Mar @ 12:54 PM

Settings > About our lab



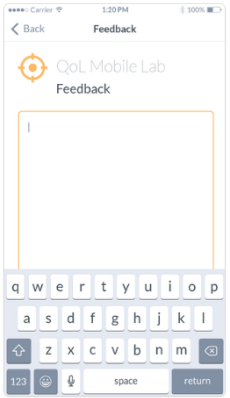
Updated 22nd Jul @ 11:59 AM

Settings > Acknowledgements



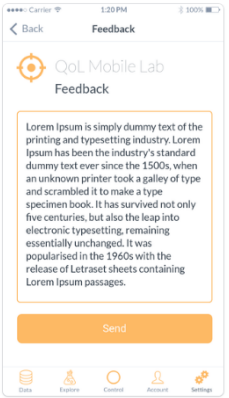
Updated 22nd Jul @ 12:00 PM

Settings > Feedback



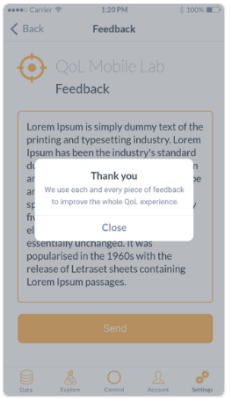
Updated 22nd Feb @ 12:04 PM

Settings > Feedback : Typing feedback



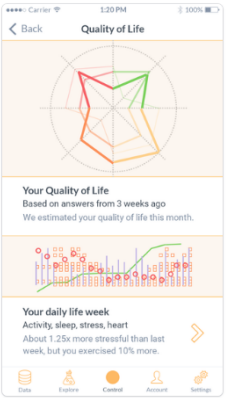
Updated 22nd Feb @ 12:04 PM

Settings > Feedback : Typed feedback



Updated 22nd Feb @ 12:04 PM

Settings > Feedback : Thank you message



Updated 22nd Jul @ 12:00 PM

Decoupled > Overview of activity and Quality of Life

Co-Calibration of Behavioural, Health, and Quality of Life Outcomes

Publication 6: Co-calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method

Vlad Manea, Katarzyna Wac. *Co-calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method*. Journal of Personalized Medicine, 10(4), MDPI, 2020. Special Issue: *PROomics: Patient Reported Outcome (PRO) and Self-Tracking for Personalized Medicine*. Impact factor 4.433, rank 10/102 (Q1) in Health Care Sciences and Services. 41p. DOI: `https://doi.org/10.3390/jpm10040203`. [55]

Abstract Inactivity, lack of sleep, and poor nutrition predispose individuals to health risks. Patient-Reported Outcomes (PROs) assess physical behaviours and psychological states but are subject of self-reporting biases. Conversely, wearables are an increasingly accurate source of behavioural Technology-Reported Outcomes (TechROs). However, the extent to which PROs and TechROs provide convergent information is unknown. We propose the coQoL PRO-TechRO co-calibration method and report its feasibility, reliability, and human factors influencing data quality. Thirty-nine seniors provided 7.4 $\pm$ 4.4 PROs for physical activity (IPAQ), social support (MSPSS), anxiety/depression (GADS), nutrition (PREDIMED, SelfMNA), memory (MFE), sleep (PSQI), Quality of Life (EQ-5D-3L), and 295 $\pm$ 238 days of TechROs (Fitbit Charge 2) along two years. We co-calibrated PROs and TechROs by Spearman rank and reported human factors guiding coQoL use. We report high PRO—TechRO correlations ($r_S \geq 0.8$) for physical activity (moderate domestic activity—light+fair active duration), social support (family help—fair activity), anxiety/depression (numeric score—sleep duration), or sleep (duration to sleep—sleep duration) at various durations (7–120 days). coQoL feasibly co-calibrates constructs within physical behaviours and psychological states in seniors. Our results can inform designs of longitudinal observations and, whenever appropriate, personalized behavioural interventions.

Keywords ambulatory assessment, physical activity, social support, anxiety, depression, nutrition, memory, sleep, health-related quality of life, wearable.

Journal of *Personalized Medicine*

MDPI

Article

Co-Calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method

Vlad Manea [1,*] and Katarzyna Wac [1,2]

[1] Quality of Life Technologies Lab, University of Copenhagen, Sigurdsgade 41, 2200 Copenhagen, Denmark; katarzyna.wac@unige.ch or wac@di.ku.dk

[2] Quality of Life Technologies Lab, University of Geneva, Route de Drize 7, 1227 Carouge, Switzerland

* Correspondence: manea@di.ku.dk

Received: 15 August 2020; Accepted: 21 October 2020; Published: 31 October 2020

Abstract: Inactivity, lack of sleep, and poor nutrition predispose individuals to health risks. Patient-Reported Outcomes (PROs) assess physical behaviours and psychological states but are subject of self-reporting biases. Conversely, wearables are an increasingly accurate source of behavioural Technology-Reported Outcomes (TechROs). However, the extent to which PROs and TechROs provide convergent information is unknown. We propose the coQoL PRO-TechRO co-calibration method and report its feasibility, reliability, and human factors influencing data quality. Thirty-nine seniors provided 7.4 $\pm$ 4.4 PROs for physical activity (IPAQ), social support (MSPSS), anxiety/depression (GADS), nutrition (PREDIMED, SelfMNA), memory (MFE), sleep (PSQI), Quality of Life (EQ-5D-3L), and 295 $\pm$ 238 days of TechROs (Fitbit Charge 2) along two years. We co-calibrated PROs and TechROs by Spearman rank and reported human factors guiding coQoL use. We report high PRO—TechRO correlations ($r_S \geq 0.8$) for physical activity (moderate domestic activity—light+fair active duration), social support (family help—fair activity), anxiety/depression (numeric score—sleep duration), or sleep (duration to sleep—sleep duration) at various durations (7–120 days). coQoL feasibly co-calibrates constructs within physical behaviours and psychological states in seniors. Our results can inform designs of longitudinal observations and, whenever appropriate, personalized behavioural interventions.

Keywords: ambulatory assessment; physical activity; social support; anxiety; depression; nutrition; memory; sleep; health-related quality of life; wearable

1. Introduction

Chronic diseases represent a significant share of the burden of disease globally [1]. They are responsible for 86% of all deaths [2]. In Europe, chronic diseases affect over 80% of adults over 65 and incur 70% of the increasing healthcare costs [3]. The most common chronic diseases are cardiovascular, pancreatic, pulmonary, and neoplastic. Unhealthy lifestyle and behaviours, such as physical inactivity, insufficient sleep, poor nutrition, and tobacco intake, explain up to 50% of the risk of chronic disease [4]. We expect the importance of the long-term risk of disease to increase as the world population is ageing [5]. As age dramatically contributes to the risk of multiple diseases [1], the healthy old is a population both inherently at risk and appropriate for primary disease prevention.

Currently, human health studies assess behaviours through a combination of self-reported outcomes [6], in particular patient-reported outcomes (*PRO*, [6]), and, more recently, patient-generated technology-reported outcomes (*TechRO*, [6]). Patient-reported outcomes include questionnaires with validated scales that assess individual outcomes momentarily or for a given recall period (e.g., *"During the past month, how often have you had trouble sleeping?"*). However, self-reports are

known to be the subject of biases related to the inherent shortcomings of participant reporting. The questionnaires are inconvenient, infrequent, memory-biased, socially conditioned, and qualitative. For example, seniors reporting physical activity tend to overestimate the amount undertaken [7], while subjective sleep is less reliable than objective sleep according to studies of sleep, ageing, and cognition [8,9].

In an attempt to address the shortcomings of self-reports and based on technological advances, we propose the *coQoL* PRO-TechRO co-calibration method. Our research primarily focuses on assessing behaviours and outcomes by combining questionnaires with devices such as smartphones and wearables, assessing multiple outcomes (e.g., physical activity, sleep, and heart rate) *momentarily*, and, if collected for a long time, also *longitudinally* [10]. Numerous studies used validated, expensive, and bulky lab-grade devices (e.g., ActiGraph), although for a limited time due to the user burden and discomfort of wearing them [11]. Conversely, consumer-friendly wearables measure continuously and objectively TechROs, increasingly more accurately, as technology progresses [12]. Also, more individuals opt for consumer-friendly wearable devices; the market size for consumer wearables will likely double by 2022 [13]. More recent research showed that consumer wearables could assess multiple behaviours accurately [14], unobtrusively [15], and continuously [16] while worn by participants during the natural unfolding of their daily lives. Overall, consumer devices are accurate and used enough to be leveraged in human health studies.

There exist prior work aiming at co-calibration of physical and psychological outcomes with technology-related ones, as discussed in this paper. We identify the previous work by following by following a semi-structured literature review detailed in Appendix A.1. Table 1 presents the PRO-TechRO co-calibration studies resulting from our literature review for the following outcomes: physical activity, social support, anxiety and depression, memory, sleep, and health-related Quality of Life. For each study, the table presents the PROs and TechROs used for co-calibration, the study design, the analysis methodology, and a summary of results. As for the PRO, the table presents the long names of the PRO instruments leveraged in the study, followed by the TechRO details, at least including the name and its form factor (consumer wearable or research-grade accelerometer, and position on the body). The study design details include its target population, sample size and age, and study duration. Past co-calibration methods range from simple descriptive statistics to inferential statistics via correlation methods, to machine learning, including regression and classification. The results bring a summary of PRO-TechRO co-calibration efforts, as presented in the paper.

To better emphasize the difference between state of the art and our work, we recall that we focus on healthy seniors and our method implies repeated sets of different PRO assessments in longitudinal daily life TechRO assessment settings, based on consumer wearables. All studies presented in Table 1 have at least one feature (marked in violet) that excludes them from co-calibrating PRO questionnaires with TechRO consumer wearables in healthy seniors *in the wild* over long periods (above the typical 7–14 days found in the literature).

Table 1 does not include studies on nutrition, since, to our best knowledge, the co-calibration of the *diet* with distant measures such as *steps* or *sleep* using questionnaire PROs and consumer wearables (or, at the very least, accelerometers) does not exist in the literature. However, there are numerous articles on energy expenditure estimates measured by consumer wearables that guide the energy intake (food types and qualities) for individuals following dietary recommendations [17–19].

As can be seen from Table 1, most studies focus on specific PROs suitable for the study aim; some of the PROs are disease-specific, which also relate to the user groups in the study (e.g., students, patients with a given condition). As for the TechROs, we observe few research-grade wearables, and many consumer-grade ones (Fitbit); mostly worn as wearable bracelets. The study design is characterized by diverse sample sizes (20–70, with very few examples of 500+ participants) and usually very short duration (7 days or less, very few beyond three weeks). We can call these co-calibration efforts momentary, as valid in these specific periods, for which the data was collected. The co-calibration method themselves used usually leverage descriptive statistical methods and correlations. The results of these co-calibrations

rarely report values $\geq$0.5. In summary, little research focused on assessing the relationships between sets of different outcomes assessed via PROs and consumer wearable TechROs in healthy seniors, in the wild, for extended periods (beyond the typical study duration of 7–14 days).

Our paper is the result of research conducted as part of the EU AAL Caregiver and ME (*CoME*, No. 14-7, 2017–2020) research project and software application. CoME aimed at self-management of health for individuals of old age at risk of mild cognitive impairments and their informal caregivers [20]. The project used numerous PROs to obtain a holistic view of the participants' health and wellbeing, by covering constructs that are both reflective (physical activity, anxiety, depression, memory, sleep) and formative (nutrition and social support) for the individual's Quality of Life (*QoL*) [21]. These constructs assess participants' health state and correspond to behavioural risk factors of dementia, as guided by the goals of the project [22–25].

Our study involved 42 seniors from Hungary and Spain. The seniors provided PROs on questionnaires chosen by the consortium of the CoME project partners along [22]. The measured outcomes included physical activity (using the International Physical Activity Questionnaire Long, or *IPAQ* [26]), social support (Multidimensional Scale of Social Support, *MSPSS* [27]), anxiety and depression (Goldberg Anxiety and Depression Scale, *GADS* [28]), nutrition (Prevention with Mediterranean Diet, *PREDIMED* [29,30] and Self-Reported Mini Nutritional Assessment, *SelfMNA* [31]), memory (Memory Failures of Everyday, *MFE* [32]), sleep (Pittsburgh Sleep Quality Index, *PSQI* [33]), and health-related Quality of Life (EuroQoL with five dimensions and three levels, *EQ-5D-3L* [34]) (Appendix B.1.1 describes the questionnaires and their validated scales in depth). Participants also provided TechROs of physical activity, sleep, and heart rate (Fitbit Charge 2 consumer wearable, [35]) during the study, for up to two years.

Our paper has three objectives. First, we aim at demonstrating the feasibility of our co-calibration method, *coQoL*, by quantifying relationships between PROs and TechROs for our sample. Second, we aim at assessing the quality of the data collected while daily life unfolded for our participants. Third, we aim at informing the design of observational (and potentially interventional) personalized behavioural studies by leveraging the results from the first two objectives.

Our paper is structured as follows. Section 1 provides an introduction. Section 2 describes our materials and methods. Section 3 foregrounds our results. Section 4 discusses our findings. Section 5 concludes the paper.

Table 1. Previous PRO-TechRO Co-Calibration Studies.

Outcome	PRO: Name	TechRO: Name, Position on Body	Study: Population, Sample, Duration	Co-Calibration: Method	Results	Reference
Physical Activity	International Physical Activity Questionnaire (IPAQ); Physical Activity for Adults Questionnaire (PAAQ)	Actical (research-grade accelerometer), right hip	Individuals, $N = 112$, age range 18–79, mean age 47, 7 days, in the wild	Spearman correlation	PAAQ and IPAQ agreed for moderate and vigorous activity ($r_S = 0.44$, $r_S = 0.2$, respectively).	Garriguet et al. (2015) [36]
Physical Activity	International Physical Activity Questionnaire (IPAQ)	Fitbit (consumer wearable), non-dominant arm; ActiGraph GT3X+ (research-grade accelerometer), right waist	Students, $N = 53$, mean age 28.10 ± 9.12, 7 days	Paired t-test, Bland Altman	No significant correlations were found between the IPAQ and the two devices.	Brewer et al. (2017) [37]
Physical Activity	Godin Leisure-Time Exercise Questionnaire (GLTEQ)	Fitbit Alta (consumer wearable), wrist	Endometrial cancer survivors, $N = 25$, mean age 62 ± 9, 30 days	U statistic	No significant correlations were found between the GLTEQ and steps.	Rossi et al. (2018) [38]
Physical Activity	International Physical Activity Questionnaire (IPAQ)	Fitbit Zip (consumer wearable), wrist; ActiGraph GTX3 (research-grade accelerometer)	Seniors, $N = 70$, age range 62–77, mean age 70.1 ± 3.3), 7 days (ActiGraph, Fitbit), 70 days (study)	Descriptive	IPAQ good for duration of activities but not intensity.	Meyer et al. (2019) [39]
Physical Activity	International Physical Activity Questionnaire (IPAQ), Patient Health Questionnaire (PHQ)	Fitbit Charge 3 (consumer wearable), wrist	Individuals with depression, $N = 8$, age range 18–95, mean age 45, 8 weeks	Descriptive	IPAQ score associated with Fitbit steps.	Santomas et al. (2020) [40]
Social Support	Pittsburgh Sleep Quality Index (PSQI), Pittsburgh Sleep Diary (PghSD), Interpersonal Support Evaluation List (ISEL), Hamilton Rating Scale for Depression (HRSD), Comorbidity Questionnaire, and others	Actiwatch 64 (accelerometer), wrist	Individuals with and without chronic insomnia, $N = 119$ (79 with insomnia), min. age 60, 7 days	Analysis of covariance (ANCOVA), ordinal logistic regression	Social support associated with lower wakefulness after sleep onset for all participants, and shorter sleep latency for those with insomnia.	Troxel et al. (2010) [41]
Social Support	Social Support Scale for Exercise Behaviour and others	ActiGraph (accelerometer)	Seniors, $N = 718$, mean age 74.4 ± 6.3, 7 days	Mixed effects regression	Socially supportive environment related to 30 min. to 1 h. of physical activity in participants with positive psycho-social attributes and up to 30 min. for those with less positive psycho-social attributes.	Carlson et al. (2012) [42]
Social support	Hospital Anxiety and Depression Score (HADS), Short Form Health Survey (SF-36)	RT3 (accelerometer), waist	Seniors, $N = 547$, mean age 79 ± 8, 7 days	Multiple regression	Number of people nearby to turn to associated with higher physical activity ($R_2 = 0.32$).	McMurdo et al. (2012) [43]
Social support	Custom questionnaire to estimate social networks and social engagement, Center for Epidemiological Studies Depression (CES-D), Montreal Cognitive Assessment (MoCA), and others	Actiwatch Spectrum (accelerometer), non-dominant wrist	Seniors, $N = 673$, mean age 71.9 ± 7.2, 3 days	Multivariate linear regression	Larger social networks ($p = 0.04$), higher network proportion of friends ($p = 0.01$), more frequent visiting with neighbors ($p < 0.01$), and more frequent attendance at organized group meetings ($p = 0.03$) associated with higher physical activity intensity levels.	Ho et al. (2018) [44]

Table 1. *Cont.*

Outcome	PRO: Name	TechRO: Name, Position on Body	Study: Population, Sample, Duration	Co-Calibration: Method	Results	Reference
Social support	Iowa-Netherlands Comparison Orientation Measure, Rochester Social Comparison Record, and others	Fitbit Flex (consumer wearable), wrist	College women, $N = 80$, mean age 20 ± 1.07, 7 days	Multilevel regression	Increase in negative social interactions (especially with friends) were consistently associated with decreases in daily physical activity with high variability.	Arigo et al. (2019) [45]
Social support	University of California Los Angeles Loneliness Questionnaire	Fitbit Flex 2 (consumer wearable), wrist	First-year college students, $N = 160$, 16 weeks (one semester)	Data mining (Apriori), machine learning classification (gradient boosting, logistic regression)	Binary level of loneliness can be detected with 80.2% accuracy. More physical activity and less sedentary behaviour associated with less loneliness.	Doryab et al. (2019) [46]
Anxiety and Depression	Patient Health Questionnaire (PHQ-9), Generalised Anxiety Disorder 7-Item Scale (GAD-7), International Physical Activity Questionnaire (IPAQ), Social Support, and others	SenseWear (accelerometer), arm	Individuals with chronic major depressive disorder or a bipolar 2 disorder, $N = 14$, age range 42–72, mean age 54.5 ± 8.7, 7 days (wear), 14 weeks (study)	Wilcoxon signed rank difference test	Physical activity results in an improvement in anxiety and depression in patients with chronic depression (median depression score decreased 38%, $p < 0.05$).	Adams et al., 2015 [47]
Anxiety and Depression	Patient Health Questionnaire (PHQ-9), Mini International Neuropsychiatric Interview (MINI), Montgomery-ÅDepression Rating Scale (MADRS)	ActiGraph GT3X+ (accelerometer)	Anxiety and depression patients, $N = 165$, age range 18–65, mean age 41.8 ± 11.6, 7 days	Analysis of variance (ANOVA), analysis of covariance (ANCOVA), paired t-tests	No significant results; depressed participants tended to be less active at light intensity ($\beta = -2.21$, $p < 0.01$).	Helgadóttir et al. (2015) [48]
Anxiety and Depression	Depression Anxiety Stress Scale (DASS)	Fitbit (consumer wearable), wrist	University students and staff, $N = 85$, mean age 22 ± 3, 3 weeks	Analysis of variance (ANOVA)	An increase in steps correlated with a decrease in depression for female participants.	Liau et al. (2018) [49]
Anxiety and Depression	University of California Los Angeles Life Stress Interview (LSI), Generalized Anxiety Disorder 7-Item Scale (GAD-7), Patient Health Questionnaire (PHQ-9)	Fitbit Charge 2 (consumer wearable), wrist	Female adolescents, $N = 30$, mean age 16.4 ± 0.8, 1 year, mean wear 7 months	Pearson correlation, Bayesian multilevel models	Within-person fluctuations in stressful life events were associated with variability in sleep duration ($r = 0.48$, $p < 0.05$). Within-person increases in sleep duration variability correlated with greater depression symptoms ($r_S = 0.38$, $p < 0.05$) while sleep regularity correlated with lesser depression ($r_S = -0.44$, $p < 0.05$).	Vidal Bustamante et al. (2020) [50]
Memory	Montreal Cognitive Assessment (MoCA), Alzheimer Disease Assessment Scale-Cognitive-Plus (ADAS-Cog Plus)	MotionWatch 8 (accelerometer), wrist	N = 151, min. age 55, mean age 71.1 ± 7.2, 5 days	Paired t-test, analysis of covariance (ANCOVA), multiple linear regression	Participants with probable mild cognitive impairment were less active and more sedentary, better ADAS-Cog Plus performance correlates with more physical activity and less sedentary behavior.	Falck et al. (2017) [51]
Memory	Self-reported learning experience (satisfaction, usefulness, and performance)	Empatica E4 (accelerometer), non-dominant wrist	College students, $N = 31$, age range 21–53, mean age 24 ± 5.9, 35 min	Machine learning (random forest, support vector machine with 3 separate kernels)	Students' perceived learning can be predicted accurately from the physiological data (89% accuracy).	Giannakos et al. (2020) [52]
Memory	Enroll-HD cognitive battery	Fitbit (consumer wearable)	Individuals with Huntington's disease, $N = 70$ (20 healthy controls), 3 uses across 8 days	Correlation tests	Medium to strong correlations between motor symptoms and cognitive tasks (r = −0.34–0.54).	McLaren et al. (2020) [53]

Table 1. *Cont.*

Outcome	PRO: Name	TechRO: Name, Position on Body	Study: Population, Sample, Duration	Co-Calibration: Method	Results	Reference
Sleep	Pittsburgh Sleep Quality Index (PSQI), Perceived Stress Scale (PSS), Short Form Health Survey (SF-12)	Q-sensor (accelerometer), dominant hand	Undergraduate students, $N = 66$, mean age 20.1 ± 1.5, 30 days	Machine learning (classification, support vector machine with 2 separate kernels)	Skin conductance, skin temperature, and acceleration classified poor/good sleep with 80–90% accuracy.	Sano et al. (2015) [54]
Sleep	Pittsburgh Sleep Quality Index (PSQI), Charlotte Attitudes Towards Sleep Scale (CATS), Sleep Hygiene Practice Scale (SHPS), and others	Fitbit Flex (consumer wearable), wrist	College students, $N = 218$, age range 18–38, mean age 20.3 ± 2.5, 7 days	Path model, Spearman correlation	Correlations between sleep duration from PSQI and Fitbit ($r_S = 0.33$, $p < 0.01$).	Peach et al. (2018) [55]
Sleep	Pittsburgh Sleep Quality Index (PSQI)	Fitbit Flex 2 (consumer wearable), wrist	Military individuals, $N = 17$, 2 weeks	Wilcoxon signed rank difference test, Spearman rank correlation test	Moderate correlation between PSQI and Fitbit sleep durations ($r_S = 0.643$, p = 0.005). Top contextual factors disrupting sleep were pain, noises, and worrying.	Thota et al. (2020) [56]
Quality of Life	Self-reported health scale (5 levels)	ActiGraph GT1M (accelerometer)	Seniors, $N = 560$, age range 65–85, mean age 71.6 ± 5.6, 7 days	Analysis of variance (ANOVA)	51% higher physical activity level was registered in those with very good health compared to those with poor and very poor health.	Lohne-Seiler et al. (2014) [57]
Quality of Life	Short Form Health Survey (SF-12), Oswestry Disability Index (ODI)	Fitbit Zip (consumer wearable)	Lumbar spine surgery patients, $N = 30$, mean age 42.6 ± 10.3, 7 days (pre-operatory wear), 6 months (post-operatory wear)	Paired t-test, Pearson correlation	No significant correlation between the improvement in steps ($p > 0.2$) or distance traveled per day ($p > 0.3$).	Mobbs et al. (2015) [58]
Quality of Life	Eastern Cooperative Oncology Group Performance Status (ECOG-PS), Karnofsky Performance Status (KPS), Patient-Reported Outcomes Measurement Information System (NIH PROMIS)	Fitbit Charge HR (consumer wearable), wrist	Advanced cancer patients, $N = 37$, age range 34–81, median age 62, 2 weeks	Spearman correlation, Kaplan-Meier curves, multivariate proportional hazards	Correlations were observed between average daily steps and ECOG-PS ($r_S = -0.63$, $p < 0.05$) and KPS ($r_S = 0.69$). Correlations were also observed between distance and ECOG-PS ($r_S = -0.61$) and KPS ($r_S = 0.66$).	Gresham et al. (2018) [59]
Quality of Life	EuroQoL with 5 Dimensions and 3 Levels (EQ-5D-3L)	Fitbit One (consumer wearable), belt	Stroke patients, $N = 27$, mean age 69.5, 7 days	Correlation tests	Quality of Life health score correlates with the number of steps (r = 0.46, $p < 0.03$).	Sasaki et al. (2018) [60]
Quality of Life	Short Form Health Survey (SF-12), Knee Injury and Osteoarthritis Outcome Score (KOOS)	Fitbit Flex (consumer wearable), non-dominant wrist	Knee arthroplasty patients, $N = 91$, mean age 67 ± 13, 7 days for 3 times points (2 weeks before surgery, day after surgery, and 2 weeks after surgery)	Multiple linear regression, Spearman rank correlation	Significant correlations of SF-12 (physical component summary) and post-operative step count ($r_S = 0.521$, $p < 0.05$).	Twiggs et al. (2018) [61]

The magenta font color highlights important limitations to the existing studies.

2. Materials and Methods

In this section, we describe the coQoL method applied within our study context (Section 2.1), participants (Section 2.2), protocol (Section 2.3), measured outcomes (Section 2.4), and data analysis (Section 2.5).

2.1. Study Context

We conducted this research as part of the EU AAL Caregiver and ME (CoME, No. 14-7), a research project and software application (2017–2020) aimed at self-management of health for individuals of old age at risk of mild cognitive impairments and their informal caregivers [20]. The goals of the CoME project were (1) to relieve the caregiver pressure through monitoring of physical, intellectual, emotional, and social wellbeing of the persons in need of care and (2) to increase seniors' wellbeing and autonomy in their environment and lower the risk of dementia [62] and healthcare costs in the long term. We achieved the goals by monitoring the seniors' state, behaviours (including physical activity and sleep), and other factors that influence the risk of dementia [22]. The study was purely observational; it did not include any behaviour intervention elements.

2.2. Study Participants

Individuals of older age, owning a smartphone or willing to use a smartphone provided to them, were invited to the care centre in their city (Spain and Hungary) to participate in the study. Forty-two individuals (mean age 69.8 $\pm$ 7.4) agreed to join CoME from January 2017 to December 2019.

2.3. Study Protocol

All individuals were informed about the study goals and gave their written informed consent for inclusion before the start of the study. We conducted the study under the Declaration of Helsinki. The institutional review board at the University of Geneva (Switzerland) approved the protocol (CoME, No. 14-7) on April 28, 2016. The study protocol pseudonymized all participant identities.

Upon the first visit at the care centre, the participants attended an informational workshop about the project aims. They received Fitbit Charge 2 wearable devices as their own (for the study duration and beyond). Furthermore, they filled a profile questionnaire and registered personal accounts in the CoME software application. Then they associated the Fitbit wearables to their accounts.

In the first and subsequent visits spread through a few months to a year from January 2017 to December 2019, the participants answered several questionnaires (PROs). Whenever needed, they were assisted by caregivers through this process. However, the participants were not explicitly informed about when they will have filled which of the questionnaires to avoid any activity pattern change before the visit.

2.4. Measured Outcomes

The study collected PROs from questionnaires with validated scales and TechROs from Fitbit Charge 2 consumer wearables. The PROs and TechROs were then co-calibrated by using the coQoL method illustrated in Figure 1.

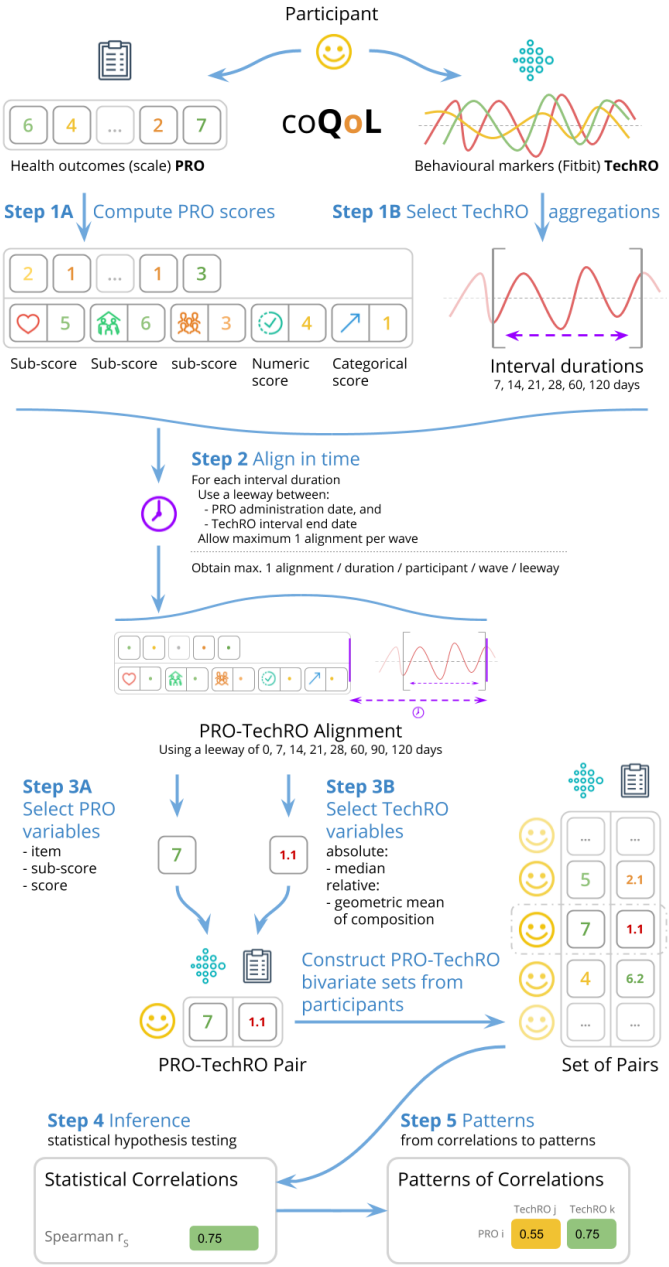


Figure 1. coQoL: a method for PRO and TechRO co-calibration (example for MSPSS PRO).

2.4.1. Patient-Reported Outcomes (Profile)

At the first visit, in the profile, participants provided their age, gender, ethnicity, profession, education, cohabitants status, height, weight, blood pressure, cholesterol, smoking, alcohol, medication (hypertension), history of personal health issues (diabetes, apnea, insomnia, hyperglycemia, stroke, infarct, depression), and history of family health issues (hypertension, diabetes, stroke, heart attack, dementia).

We included in the analysis participants who self-reported mild disease. We selected participants into three *health groups*: (1) all participants (denoted as the *all* health group), (2) only the healthy participants (*healthy*), and (3) only those with mild disease (*diseased*).

2.4.2. Patient-Reported Outcomes (PROs)

During several study visits, the participants provided answers to questionnaires for eight PROs: physical activity (IPAQ), social support (MSPSS), anxiety and depression (GADS), nutritional adherence to the Mediterranean diet (PREDIMED), nutrition (SelfMNA), memory (MFE), sleep (PSQI), and health-related QoL (EQ-5D-3L). Appendix B.1.1 describes the questionnaires in depth.

We administered the questionnaires in the languages of the respondents (Spanish or Hungarian). Appendix B.1.2 elaborates on the administration of the questionnaires.

The days of administration resulted in distinct periods of answers separated by a few months to one year. We denote these periods as *waves* of participation.

We coded the answers and computed the *scores* (and *sub-scores*, where available) according to the validated scale of each questionnaire. This procedure is depicted as Step 1A in Figure 1. Appendix B.1.3 provides details on the scoring.

We derived for the analysis the following PRO-based *variables*: (1) the individual questions in the questionnaire (denoted *items*), the *sub-scores* (where available), and the *scores* (where available). Most scales have a *numeric score* and a *categorical score*. Most sub-scores are numeric.

This procedure corresponds to Step 3A in Figure 1. All variables can be seen in Table 2. Appendix B.1.4 details the variable derivation for PROs.

Table 2. Variables derived from the PROs.

Outcome	Scale	Item Variables	Score Variables	Total
Physical Activity	International Physical Activity Questionnaire (IPAQ) [26]	15: 11 for the combinations of domains and intensities, 4 for the domain totals	8: 4 for the domain numeric scores, 3 for the intensity numeric scores, and 1 for the overall numeric score	23
Social Support	Multi-Dimensional Scale Perceived Social Support (MSPSS) [27]	12 for the items	5: 3 for the numeric sub-scores and 2 for the numeric and categorical scores	17
Anxiety and Depression	Goldberg depression and anxiety scale (GADS) [28]	18 for the items	2 for the numeric and categorical scores	20
Nutrition Mediterranean	Prevention with Mediterranean Diet (PREDIMED) [29,30]	14 for the items	2 for the numeric and categorical scores	16
Nutrition	Self-Reported Mini Nutritional Assessment (SelfMNA) [31]	5 for the items	2 for the numeric and categorical scores	7
Memory	Memory Failures of Everyday (MFE) [32]	28 for the items	2 for the numeric and categorical scores	30
Sleep	Pittsburgh Sleep Quality Index (PSQI) [33]	18 for the items	10: 8 for the sub-scores and 2 for the numeric and categorical scores	28
Health-Related Quality of Life	EuroQoL health questionnaire (EQ-5D-3L) [34]	6 for the items	0 (the scores coincide with the items)	6

2.4.3. Technology-Reported Outcomes (TechROs)

We collected the behavioural wearable markers from the daily aggregates provided by the Fitbit daily activity summary application programmable interface (API) [63]. Appendix B.2.1 motivates our choice for Fitbit as a personal wearable activity monitor in the context of our study.

We processed the wearable data by aggregating it over consecutive days in *aggregate intervals* spanning from 7 to 120 days. We included in the analysis only days with at least 21 hours of Fitbit measurement as *valid days*. Then we required each aggregate interval to have at least 70% valid days. This procedure corresponds to Step 1B in Figure 1. Appendix B.2.2 details the data processing.

The Fitbit consumer wearables provided TechROs as *raw* (energy expenditure, steps, heart rate) and *processed* according to Fitbit's internal activity recognition algorithms (sedentary duration, durations of physical activity at the light, fair, and vigorous intensities, and sleep) [35].

We derived TechRO-based *variables* in two *amounts*. The *absolute* amount refers to the TechROs enumerated above. For this amount, we computed for each interval the median of daily measurements.

We derived the *relative* amount variables from the total daily durations of physical activity (and, separately, physical activity and sleep for all 24 h [64]), transformed into compositions [65], and expressed as centred log-ratios (*CLR*). For this amount, we computed for each interval the geometric mean of the daily compositions.

Each amount has two *families*. The absolute amount has the *(absolute) raw* family (for *energy* expenditure, *steps*, and *heart rate*) and the *(absolute) processed* family (for the durations of *sleep* and physical activity at the four intensities reported by Fitbit: *sedentary*, *light*, *fair*, and *vigorous*). As Fitbit had not provided thresholds for the reported physical activity intensities (see [66–68]), we also included cumulative variables of adjacent pairs of intensities, e.g., *light+fair*. Furthermore, we included a total daily *active* duration that added all non-sedentary intensity durations.

The relative amount has the *(relative) centred log-ratio for physical activity* family (*CLR PA*) that adds for each day the durations of physical activity at the four intensities above, and the *(relative) centred log-ratio for physical activity and sleep* family (*CLR PA+S*) that adds for each day the durations of physical activity (four intensities) and sleep.

This procedure corresponds to Step 3B in Figure 1. All variables can be seen in Table 3. Appendix B.2.3 provides details on the variable derivation for TechROs.

Table 3. Variables derived from the TechROs.

Amount	Family	Outcome	Variable	Unit
Absolute	Raw	Energy	Median count over 7 days Median count over 14 days Median count over 21 days Median count over 28 days Median count over 60 days Median count over 90 days Median count over 120 days	kcal.
		Steps	Median count over [...] days	count
		Heart rate	Median beats over [...] days	bpm.
	Processed	Sedentary	Median duration over [...] days	min.
		Sedentary+Light	Median duration over [...] days	
		Light	Median duration over [...] days	
		Light+Fair	Median duration over [...] days	
		Fair	Median duration over [...] days	
		Fair+Vigorous	Median duration over [...] days	
		Vigorous	Median duration over [...] days	
		Active	Median duration over [...] days	
		Sleep	Median duration over [...] days	
Relative	CLR PA	Sedentary	Geometric mean over [...] days	-
		Light	Geometric mean over [...] days	
		Fair	Geometric mean over [...] days	
		Vigorous	Geometric mean over [...] days	
	CLR PA+S	Sedentary	Geometric mean over [...] days	
		Light	Geometric mean over [...] days	
		Fair	Geometric mean over [...] days	
		Vigorous	Geometric mean over [...] days	
		Sleep	Geometric mean over [...] days	

2.4.4. Co-Calibration (PROs vs. TechROs)

We co-calibrated PROs with TechROs by alignment. Concretely, for a PRO variable to align to a TechRO variable, the administration date of the former must have been within a set duration (0–120 days) from the end date of the latter.

To account for small samples, we allowed a *leeway* (0–120 days) between the end of the TechRO monitoring interval and the PRO scale administration date.

For each participant, we included only the last alignment in a wave, to discard repeated answers within a few minutes and reduce bias towards overly diligent responders.

When we aligned PROs with TechROs of increasing durations, the number of paired observations decreased; we thus required a minimum of 10 observations to have a nontrivial size [69].

For each PRO-TechRO pair, we reported the highest correlation among all *aggregation intervals* of TechRO (7–120 days) aligned to match the PRO administration date. We included only *significant* correlations, i.e., those correlation coefficients whose 95% confidence interval maintained sign. This procedure corresponds to Step 2 in Figure 1. Appendix B.3 elaborates on the details of the PRO-TechRO variable alignment.

2.5. Data Analysis

We conducted descriptive and inferential analyses of the PROs and TechROs. We then analyzed patterns from the analyses.

2.5.1. Descriptive Analysis (PROs and TechROs)

The descriptive analysis consisted of summary statistics (median, mean, and standard deviation, or *SD*) based on *groups* of participant-wave characteristics. In our study, we analyzed the participants by their *health*, *country*, and *gender* self-reported groups. For PROs, we observed the statistics across waves. Appendix B.1 elaborates on the analysis of the PRO variables. For TechROs, we observed the statistics across the entire study period and by counting valid days, described in depth in Appendix B.2. Appendix B.3.1 details the descriptive analysis procedure.

2.5.2. Inferential Analysis (PROs vs. TechROs)

We co-calibrated PRO variables with TechRO variables by applying the Spearman [70] statistical test on each pair of PRO-TechRO variables resulting from the alignments. The Spearman r_S statistical correlation coefficient measures the direction and strength of the association between two variables. We used the SciPy library [71] to implement the Spearman correlations. Appendix B.3.2 elaborates on the motivation and assumptions for the inferential analysis. This procedure corresponds to Step 4 in Figure 1.

2.5.3. Pattern Analysis (PROs vs. TechROs)

We used the results from the inferential analysis to highlight informative PRO variables and pairs of PRO-TechRO. This procedure corresponds to Step 5 in Figure 1. We employed two metrics that focus on the number of correlations (a high number of significant correlations with TechRO variables indicates that the PRO variable is informative) and the quality of the correlations (where possible, a strong significant correlation with other significant correlations in its vicinity indicates that the PRO-TechRO correlation is informative).

The first metric, denoted *total*, counts all strong correlations ($r_S \geq 0.5$) for a given PRO variable and highlights those PRO variables that correlate with the most TechRO variables. We applied this metric to all PRO variables.

The second metric, denoted *contour*, can only apply for variables that can be ordered by a criterion. For our study, we ordered TechRO physical activity variables by their intensities (from *sedentary* to *vigorous*). We applied this metric on strong and significant correlations ($r_S \geq 0.8$) between a PRO

and a TechRO physical activity intensity variable. The metric counted the maximum number of adjacent significant correlations of the same PRO variable (at *lower* and, separately, *higher* intensities) such that they would form a contiguous sequence of significant correlations that maintained the sign. Appendix B.3.3 further explains and exemplifies this metric.

3. Results

In this section, we report the results from the study participants (Section 3.1) and analyses (descriptive in Section 3.2, inferential in Section 3.3, and patterns in Section 3.4) as well as two use case examples for coQoL (Section 3.5).

3.1. Study Participants

Forty-two seniors (mean age 69.8 ± 7.4) signed up for the study. From these, 39 participants (mean age 70.0 ± 7.2, 22 women, 26 from Spain 26 and 13 from Hungary) provided at least one PRO; three participants were disqualified. Out of the qualified participants, 28 reported no health condition (thus being in the *healthy* health group) and 11 reported a mild health condition (forming the *diseased* health group). Participant characteristics are available in Table 4.

Table 4. Characteristics of Study Participants.

Variables	Mean (SD) or n [%]		Variables	Mean (SD) or n [%]	
	Spain	Hungary		Spain	Hungary
Count	26 [66.7%]	13 [33.3%]	Health status		
Age	69.2 (±5.7)	71.5 (±9.1)	Healthy	18 [46.2%]	10 [25.6%]
Gender			Diseased	8 [20.5%]	3 [7.7%]
Women	15 [38.5%]	7 [17.9%]	Smoking		
Men	11 [28.2%]	6 [15.4%]	Yes	5 [12.8%]	1 [2.6%]
Education			No	21 [53.8%]	12 [30.8%]
Primary	7 [17.9%]	0 [0.0%]	Alcohol		
Secondary	5 [12.8%]	3 [7.7%]	Never	10 [25.6%]	4 [10.3%]
High school	5 [12.8%]	1 [2.6%]	Monthly	5 [12.8%]	5 [12.8%]
University	9 [23.1%]	9 [23.1%]	Weekly	7 [17.9%]	1 [2.6%]
Living			Few days	1 [2.6%]	2 [5.1%]
Alone	11 [28.2%]	3 [7.7%]	Daily	3 [7.7%]	1 [2.6%]
+Partner	14 [35.9%]	10 [25.6%]	Systolic blood pressure	146.2 (±63.2)	124.7 (±15.0)
+Children	1 [2.6%]	0 [0.0%]	Body mass index	25.5 (±4.64)	28.5 (±4.1)

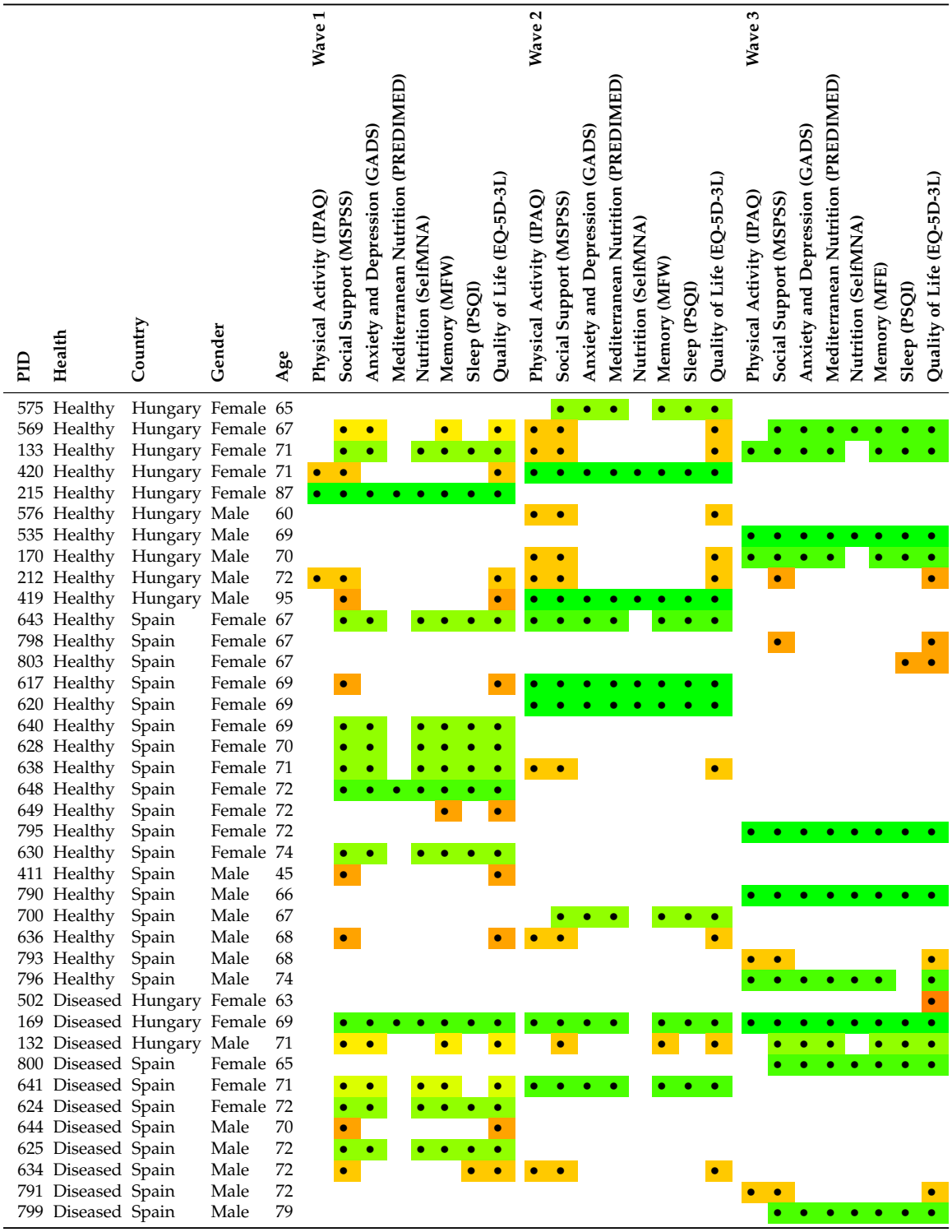
+: addition to the previous row.

3.2. Descriptive Analysis (PROs and TechROs)

3.2.1. Patient-Reported Outcomes (Questionnaires)

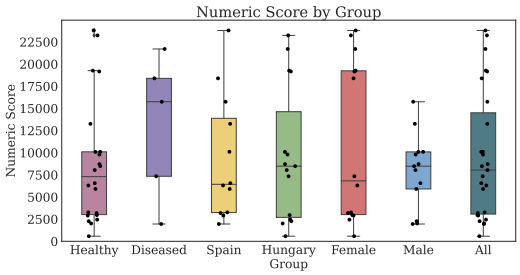
Three waves of PRO participation resulted from January 2017 to December 2019: wave 1 (mid-2018), wave 2 (end-2018 and start-2019), and wave 3 (mid-2019). Table 5 illustrates the waves of participation for each participant and questionnaire.

Table 5. PRO count answers by wave and questionnaire ($N = 39$ participants).

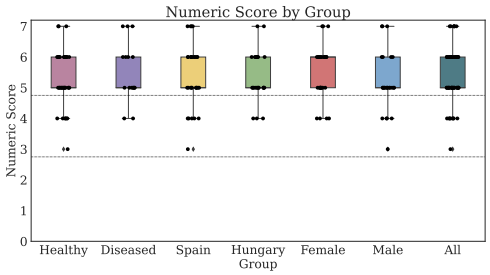


Color coding: from orange (fewer scales answered in a wave) to yellow to green (more answered).

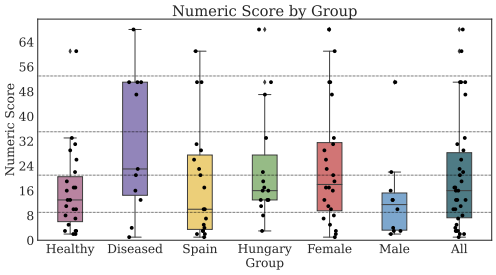
Figures 2 and 3 depict the numeric scores for all patient-reported outcome scales. Appendix B.1 details the results in-depth for each PRO variable.



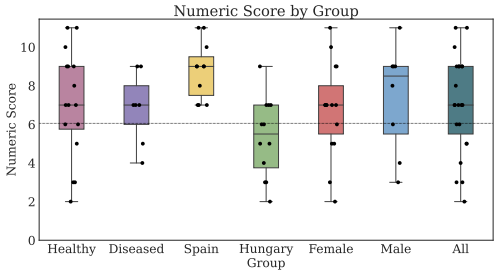
(a) Physical Activity (IPAQ): higher score ≈ more physical activity



(b) Social Support (MSPSS): higher score ≈ more social support

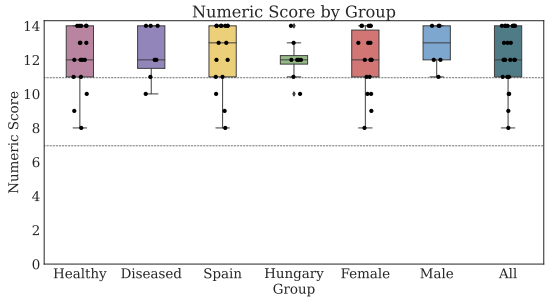


(c) Anxiety and Depression (GADS): higher score ≈ more anxiety / depression

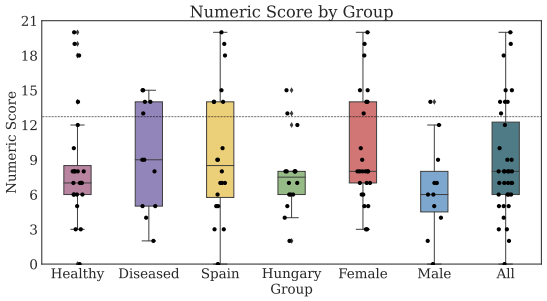


(d) Mediterranean Nutrition (PREDIMED): higher score ≈ more adherence

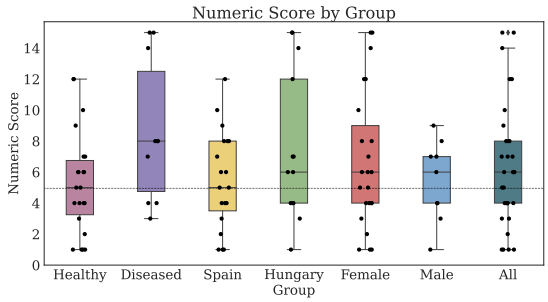
Figure 2. Numeric scores for Physical Activity, Social Support, Anxiety and Depression, and Mediterranean Nutrition. Dotted markings delimit levels of the categorical score, where available (1 of 2).



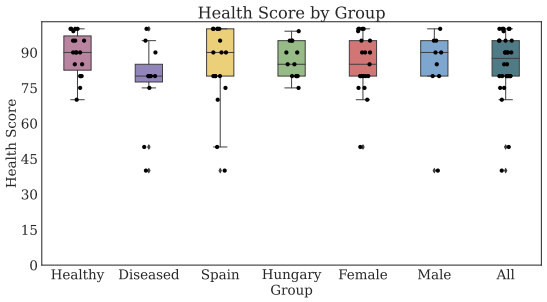
(a) Nutrition (SelfMNA): higher score ≈ less chances of malnutrition



(b) Memory (MFE): higher score ≈ more chances of memory failure



(c) Sleep (PSQI): higher score ≈ lower sleep quality



(d) Health-Related Quality of Life (EQ-5D-3L): higher score ≈ better health

Figure 3. Numeric scores for Nutrition, Memory, Sleep, and Health-Related Quality of Life. Dotted markings delimit levels of the categorical score, where available (2 of 2).

3.2.2. Technology-Reported Outcomes (Fitbit)

Thirty-two participants provided both PROs and TechROs. Figures 4 and 5 depict the counts of participants by monitored and valid Fitbit days, respectively. Figures 6 and 7 depict the distribution of monitored and valid Fitbit days, respectively. Figures 8 and 9 depict the medians of TechROs across the entire monitoring period for the participants. Appendix B.2 provides additional details on compliance and analyses each TechRO in-depth.

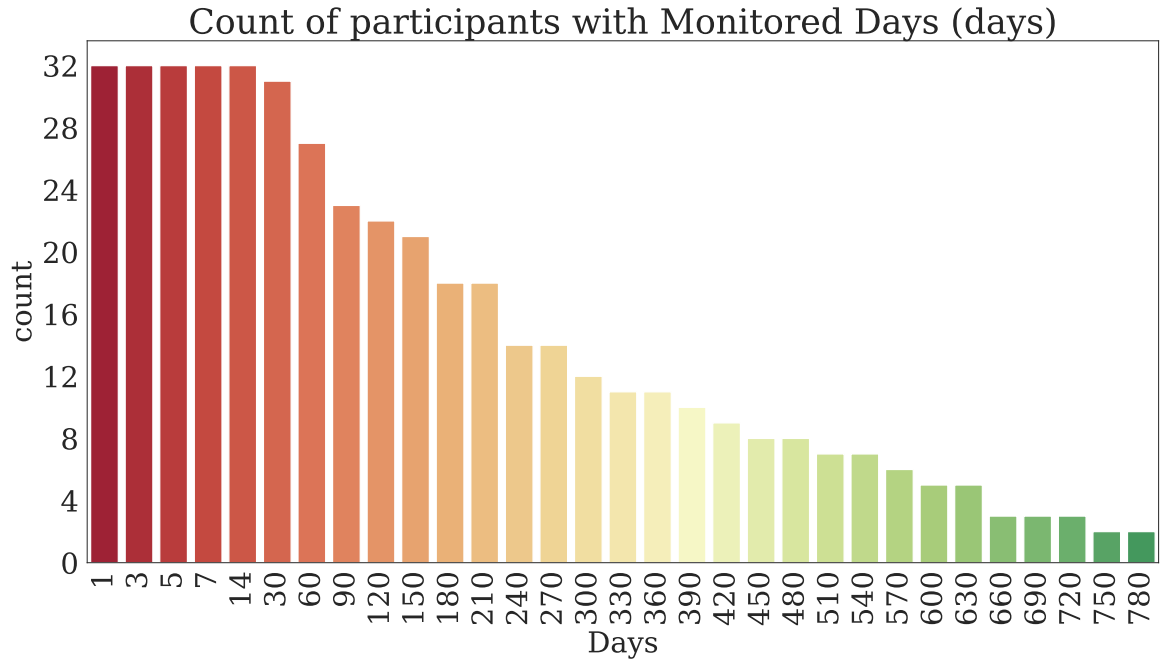


Figure 4. Count of seniors with at least the given monitored days of Fitbit (TechRO).

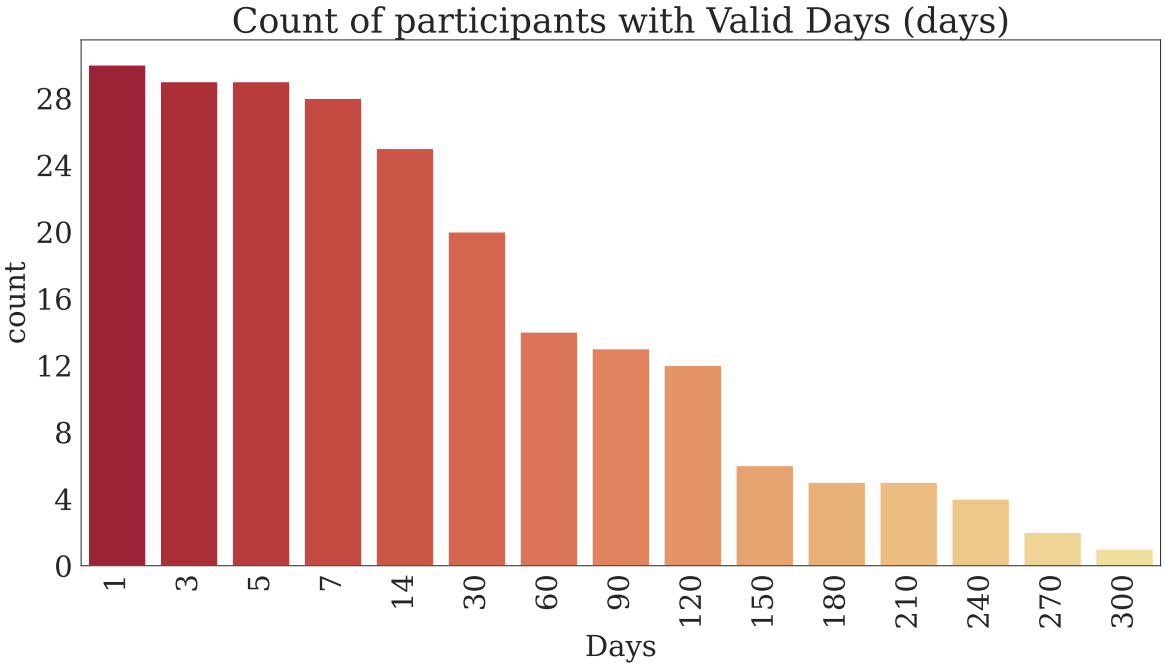


Figure 5. Count of seniors with at least the given valid days of Fitbit (TechRO).

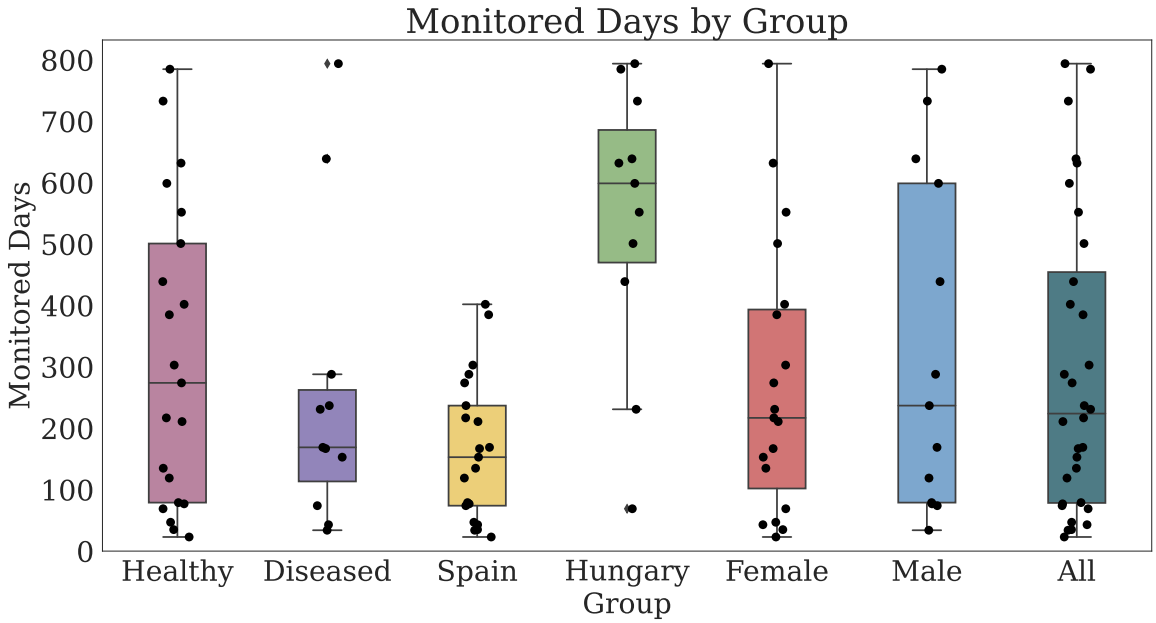


Figure 6. Days of Fitbit (TechRO) monitored days for seniors with at least one PRO.

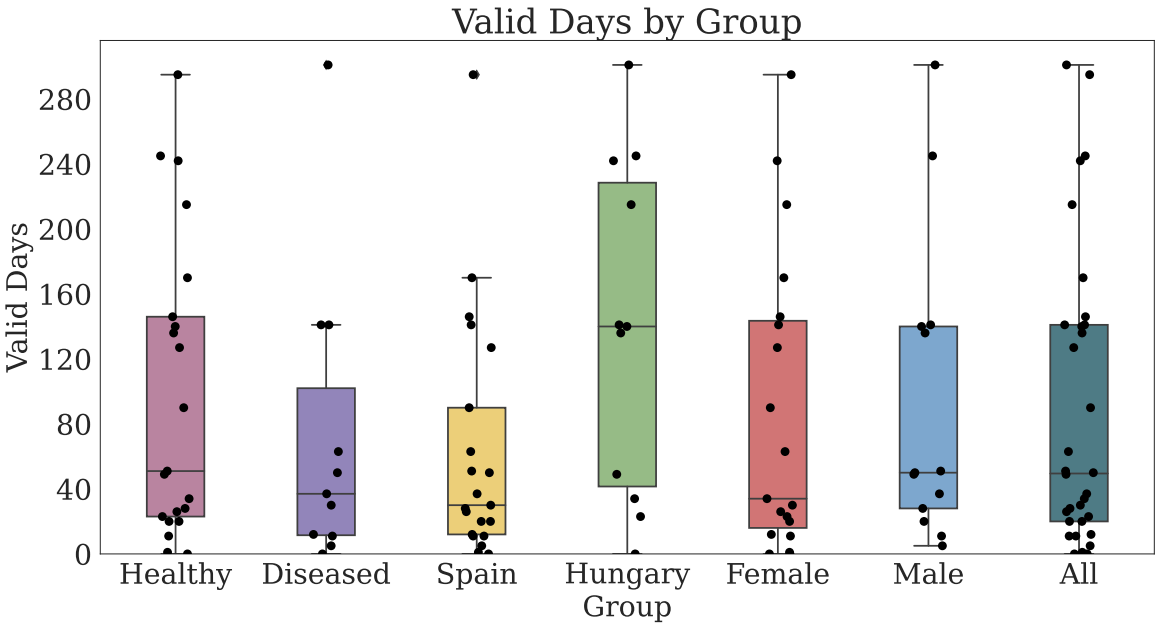
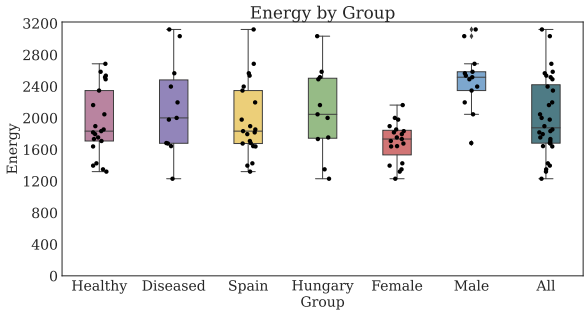
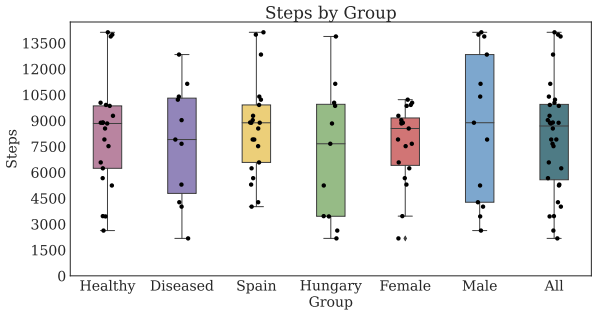


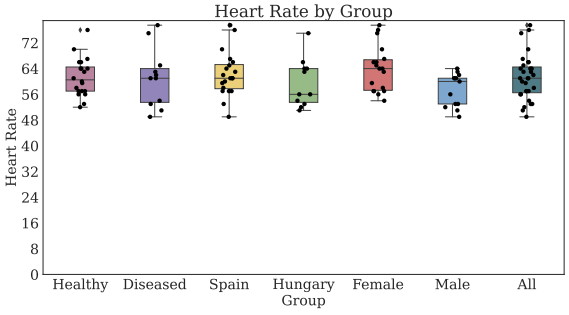
Figure 7. Days of Fitbit (TechRO) valid days data for seniors with at least one PRO.



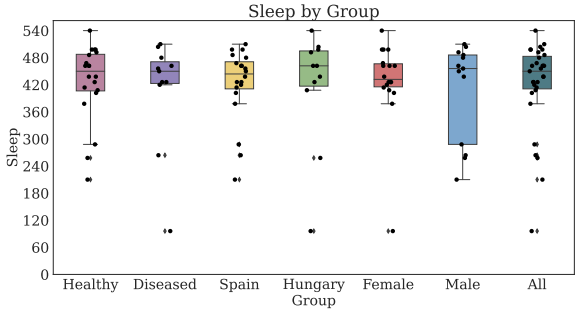
(a) Energy (kcal.)



(b) Steps (count)

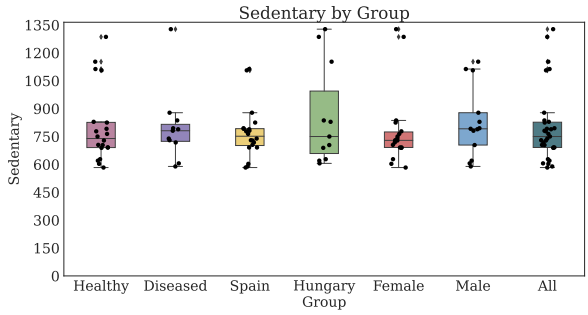


(c) Heart Rate (bpm.)

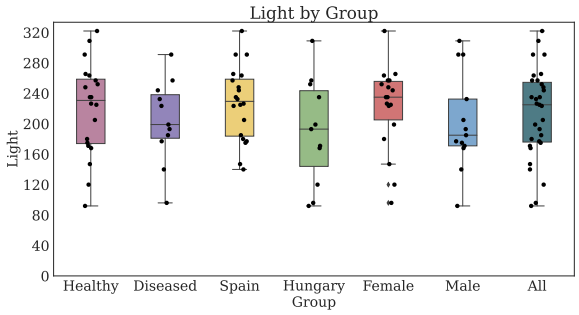


(d) Sleep Duration (min.)

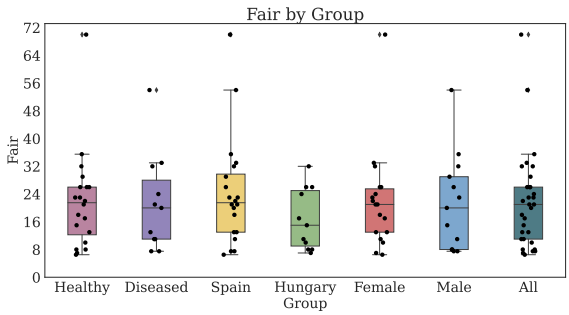
Figure 8. Median values of TechROs (Fitbit) across the entire monitoring period: energy, steps, heart rate, and sleep (1 of 2).



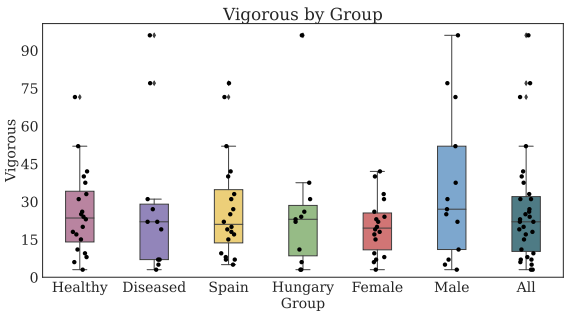
(a) Sedentary Duration (min.)



(b) Light Duration (min.)



(c) Fair Duration (min.)



(d) Vigorous Duration (min.)

Figure 9. Median values of TechROs (Fitbit) across the entire monitoring period: physical activity (2 of 2).

3.3. Inferential Analysis (PROs vs. TechROs)

Appendix C.2 elaborates on the Spearman rank correlations resulted from the inferential analysis on each questionnaire and PRO-TechRO variable pair.

3.4. Pattern Analysis (PROs vs. TechROs)

We report further the results of the pattern analysis for each questionnaire: physical activity (Section 3.4.1), social support (Section 3.4.2), anxiety and depression (Section 3.4.3), Mediterranean nutrition (Section 3.4.4), nutrition (Section 3.4.5), memory (Section 3.4.6), sleep (Section 3.4.7), and health-related Quality of Life (Section 3.4.8).

3.4.1. coQoL for Physical Activity (IPAQ vs. Fitbit)

We report the correlations of PRO physical activity variables (IPAQ) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Physical Activity Outcomes by Total Numbers of Correlations

Table 6 highlights the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all TechRO families by health group.

Table 6. PROs with high total count of significant Spearman correlations ($r_S \geq 0.5$) with TechROs.

Outcome	PRO	Health	PRO	TechRO Families				
			Item/Sub-Score/Score	Raw	Processed	CLR PA	CLR PA+S	All
Physical activity	IPAQ	All	Domestic moderate activity		4	2	2	8
Physical activity	IPAQ	All	Domestic+garden total activiy		3	2	3	8
Physical activity	IPAQ	All	Garden moderate activity		4	2	1	7
Physical activity	IPAQ	All	Leisure moderate activity	1	3	2	1	7
Physical activity	IPAQ	Healthy	Domestic moderate activity	2	4	3	2	11
Physical activity	IPAQ	Healthy	Garden moderate activity		6	4		10
Physical activity	IPAQ	Diseased	Garden vigorous activity	1	6	3	2	12
Physical activity	IPAQ	Diseased	Leisure vigorous activity	2	6	2	2	12
Physical activity	IPAQ	Diseased	Work vigorous activity	1	5	3	2	11
Physical activity	IPAQ	Diseased	Work moderate activity	2	5	1	2	10
Social support	MSPSS	All	Q8: family talks about problems		4	3	3	10
Social support	MSPSS	All	Q11: family willing to help make decisions	1	5	2	2	10
Social support	MSPSS	Healthy	Q3: family tries to help	1	6	3	4	14
Social support	MSPSS	Healthy	Q6: friends try to help	1	7	2	4	14
Social support	MSPSS	Healthy	Q9: friends share joys and sorrows	1	6	2	4	13
Social support	MSPSS	Healthy	Q12: friends talk problems	1	7	2	3	13
Social support	MSPSS	Healthy	Q10: special person cares about feelings		7	1	4	12
Social support	MSPSS	Healthy	Friends numeric sub-score	1	6	2	3	12
Social support	MSPSS	Diseased	Q2: special person shares joys and sorrows	1	5			6
Social support	MSPSS	Diseased	Significant other numeric sub-score	1	4		1	6
Anxiety and depression	GADS	All	Q6D: lost weight due to poor appetite		5	3	4	12
Anxiety and depression	GADS	All	Q8A: worried about own health		4	4	2	10
Anxiety and depression	GADS	All	Q1D: lacking energy		3	3	4	10
Anxiety and depression	GADS	Healthy	Q2D: lost interest in things		6	3	3	12
Anxiety and depression	GADS	Diseased	Q2A: worrying a lot	2	6	2	1	11
Mediterranean nutrition	PREDIMED	All	Categorical score	2	4	3	1	10
Mediterranean nutrition	PREDIMED	All	Numeric score	1	3	4	1	9
Mediterranean nutrition	PREDIMED	All	Q12: nuts use	2	2	1	2	7
Mediterranean nutrition	PREDIMED	All	Q14: sofrito use		2		5	7
Mediterranean nutrition	PREDIMED	Healthy	Q4: fruit use	1	3	2	1	7
Mediterranean nutrition	PREDIMED	Healthy	Categorical score		2	2	2	6
Nutrition	SelfMNA	All	Categorical score		2	2	2	6
Nutrition	SelfMNA	Healthy	Categorical score		1	2	2	5
Nutrition	SelfMNA	Diseased	Q2: weight lost	1	3	1	2	7
Nutrition	SelfMNA	Diseased	Q1: food intake declined	1	2	1	2	6

Table 6. *Cont.*

Outcome	PRO	Health	PRO Item/Sub-Score/Score	TechRO Families				
				Raw	Processed	CLR PA	CLR PA+S	All
Memory	MFE	All	Q12: having difficulty picking up a new skill		6	1	4	11
Memory	MFE	All	Q14: forgetting to do planned things		5	2	3	10
Memory	MFE	All	Q6: forgetting time of events		4	3	2	9
Memory	MFE	Healthy	Q6: forgetting time of events	1	7	3	3	14
Memory	MFE	Healthy	Q15: forgetting details of done things		7	2	4	13
Memory	MFE	Healthy	Q12: having difficulty picking up a new skill		6	3	3	12
Memory	MFE	Healthy	Q14: forgetting to do planned things	1	6	2	3	12
Memory	MFE	Diseased	Q13: having a word on the tip of the tongue	1	7	3	2	13
Memory	MFE	Diseased	Q25: getting lost in often visited place		7	3	2	12
Sleep	PSQI	All	Q7: trouble staying awake driving, eating, socializing	2	5	4	3	14
Sleep	PSQI	All	Q4: duration of actual sleep	1	5	3	2	11
Sleep	PSQI	All	Daily dysfunction numeric sub-score	1	4	3	2	10
Sleep	PSQI	Healthy	Q4: duration of actual sleep	1	5	3	2	11
Sleep	PSQI	Healthy	Q5C: trouble sleeping due to using the bathroom		4	4	2	10
Sleep	PSQI	Healthy	Q7: trouble staying awake driving, eating, socializing	2	5	3		10
Sleep	PSQI	Healthy	Daily dysfunction numeric sub-score	2	3	3	1	9
Sleep	PSQI	Diseased	Daily dysfunction numeric sub-score	2	4	1		7
Sleep	PSQI	Diseased	Q6: duration of actual sleep		4	2		6
Quality of Life	EQ-5D-3L	All	Q6: health state today		4	1	3	8
Quality of Life	EQ-5D-3L	All	Q4: pain/discomfort		2	1	3	6
Quality of Life	EQ-5D-3L	Healthy	Q4: pain/discomfort		4	2	1	7
Quality of Life	EQ-5D-3L	Diseased	Q5: anxiety/depression	2	3			5

Color coding: from orange (less correlations) to green (more correlations).

In the health group with all participants, when assessing totals of correlations, PRO *moderate* activity in the *domestic*, *garden*, and *leisure* domains correlated with the most TechROs (Table 6).

In the group with healthy participants, PRO *moderate activity* in the *domestic* and *garden* domains had the most correlations with TechROs as well. The *domestic moderate* and *garden moderate* activity were also the only two PROs highlighted by the total metric in the groups with all and healthy participants.

In the group with diseased participants, PRO *vigorous* in the *garden* and *leisure* domains correlated with the most TechROs, followed by the PRO *moderate* and *vigorous* activities in the *work* domain (Table 6).

Physical Activity Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (IPAQ) and TechRO variables (Fitbit) in Table 7.

In the health group with all participants, when assessing strong correlations, the PRO *domestic moderate* activity had a small contour of correlations with the TechRO *light+fair* physical activity. Also, the PRO *work vigorous* activity may explain the TechRO *active* duration without a contour (Table 7, rows with Health: All).

In the group with healthy participants, only two strong correlations emerged without contours. PRO work moderate and total activity correlated with the TechRO fair activity duration (Table 7, rows with Health: Healthy).

In the group with diseased participants, we found numerous correlations with and without contours in the *work* domain. A positive relationship with a broad contour occurred between PRO *work moderate* activity and TechRO *fair* activity duration. Furthermore, PRO *work moderate* activity correlated negatively with TechRO *sedentary* duration. However, work activity at the two extreme intensities (*walking* and *vigorous*) also correlated negatively with relative *light* activity (Table 7, rows with Health: Diseased and PRO Domain: Work).

For the PRO *garden* domain, PRO *vigorous* activity correlated negatively with contours with TechRO relative *sedentary* and *light* activity, indicating that it may redistribute physical activity across the other intensities over the day (Table 7, rows with Health: Diseased and PRO Domain: Garden).

For the PRO *leisure* domain, *walking* activity correlated without contours with *energy* and *steps*. PRO *leisure vigorous* activity correlated positively with TechRO *fair+vigorous* activity durations and negatively with TechRO absolute *sedentary* and relative *light* durations. The PRO *leisure total* activity had a correlation with contour consistent with the previous correlation: negative relationship with TechRO *sedentary+light* activity (Table 7, rows with Health: Diseased and PRO Domain: Leisure).

The PRO *vigorous* activity in the *work* domain appeared in both groups with all and diseased participants. However, its correlations were divergent: for all participants, the *work vigorous* associated with the total daily *activity*, while for the mildly diseased, it may replace *light* activity. The *moderate activity* at *work* had inverse relations with *fair* activity for diseased (positive) and healthy (negative) participants. However, for the diseased, the correlation had a broad contour, while for the healthy it had none. In this case, the latter relation may have been a false positive (Table 7, rows with PRO Domain: Work).

Across numerous PROs, the TechRO of *sedentary* activity correlated strongly only for diseased participants and mostly in relative families. PRO *moderate* to *vigorous* activity at *work*, in the *garden*, and for *leisure* all negatively correlated with TechRO daily *sedentary* duration. These results indicate that *moderate* activity may contribute to lower measured TechRO *sedentary* duration, but the redistributions of daily time to other TechRO intensities may vary between TechRO *fair* and *vigorous* intensities. (Table 7, rows with Health: Diseased and TechRO Variable: Sedentary).

Table 7. Summary of strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of physical activity (IPAQ scale) and TechROs (Fitbit wearable).

	PRO		TechRO			Correlation/Contour			
Health	**Domain**	**Variable**	**Amount**	**Family**	**Variable**	**Lower**	r_S	**Higher**	
All	Work	Vigorous activity	Absolute	Processed	Active		+0.8		
All	Domestic	Moderate activity	Absolute	Processed	Light+fair	+0.7	+0.8	×	
Healthy	Work	Moderate activity	Absolute	Processed	Fair	×	−0.8	×	
Healthy	Work	Total activity	Absolute	Processed	Fair	×	−0.8	×	
Diseased	Work	Walking activity	Relative	CLR PA	Light	−0.7	−0.8	×	
Diseased	Work	Moderate activity	Absolute	Processed	Fair	×	+0.8	+0.7	+0.7
Diseased	Work	Moderate activity	Relative	CLR PA	Sedentary		−0.8	×	
Diseased	Work	Vigorous activity	Relative	CLR PA	Light	−0.7	−0.8	−0.6	
Diseased	Garden	Vigorous activity	Relative	CLR PA	Light	−0.7	−0.8	−0.5	
Diseased	Garden	Vigorous activity	Relative	CLR PA+S	Sedentary		−0.8	−0.7	
Diseased	Leisure	Walking activity	Absolute	Raw	Energy		+0.8		
Diseased	Leisure	Walking activity	Absolute	Raw	Steps		+0.8		
Diseased	Leisure	Vigorous activity	Absolute	Processed	Fair+Vigorous	×	+0.8	+0.6	
Diseased	Leisure	Vigorous activity	Relative	CLR PA	Sedentary		-0.8	×	
Diseased	Leisure	Vigorous activity	Relative	CLR PA	Vigorous	×	+0.8		
Diseased	Leisure	Vigorous activity	Relative	CLR PA+S	Light	−0.7	−0.8	×	
Diseased	Leisure	Total activity	Absolute	Processed	Sedentary+light	−0.6	−0.8	×	

Color coding: from orange (weak correlation) to green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

Physical Activity Outcomes Highlighted by Both Metrics

For the health group with all participants, the *domestic moderate* activity appeared with both metrics. This result is in concordance with the strong correlations in the PRO *domestic* domain mentioned above (Tables 6 and 7, rows with Health: All).

In the group with diseased participants, the total metric results confirmed those using the contour metric for the PRO *work* domain at *moderate* and *vigorous* intensities (Tables 6 and 7, rows with Health: Diseased).

Physical Activity Outcomes Interpretation

In the health group with all participants, we observed several "expected" correlations. The PRO *domestic moderate* activity associated with the TechRO absolute *light+fair* activity duration. This effect is only visible for the total metric, indicating that PRO *domestic* and *garden moderate* activity may redistribute physical activity across numerous TechRO intensities.

In the group with diseased participants, PRO *work moderate* associated with the TechRO absolute *fair* activity duration. For the same health group, *leisure walking* activity correlated with both *energy* and *steps*, while PRO *vigorous* activity correlated with both absolute *fair+vigorous* activity and relative *vigorous* activity (when including sleep).

In this group, we also found "expected" correlations between PROs and TechRO *sedentary* duration. PRO *moderate* activity at *work*, *vigorous* activity in the *garden*, and *vigorous* activity for *leisure* associated negatively with TechRO *sedentary* duration. The TechRO *sedentary+light* duration associated negatively with the PRO *total active* effort as well.

Other associations indicate potential activity replacements (within TechRO) for the same health group (diseased). Walking at *work* associated negatively with the relative duration of activity at the *light* intensity, indicating that, when they *walk* at *work*, they tend to perform less *light* activity elsewhere. Also, the *vigorous* activity effort may replace *light* activity duration during the day, indicating that the participants tend to limit their physical activity to a narrow spectrum of intensities.

The distribution of results per families of TechROs indicates that for the groups with all participants and the healthy, the absolute families may provide most, if not all, strong correlations. However, for the diseased group, measuring the entire physical activity duration and including sleep uncovered associations weaker or non-significant otherwise. For this group, measuring only raw *energy* or *steps* TechROs may be indicative of their *leisure walking* efforts, potentially useful for more sedentary participants who do not work.

Both metrics highlighted all IPAQ domains except *transport*. The PRO *transport* physical activity was not indicative of TechRO physical activity measures, potentially due to the lower and fewer correlations with *transport*. However, the raw responses indicate that *transport walking* activity may associate with the *numeric score* of physical activity.

3.4.2. coQoL for Social Support (MSPSS vs. Fitbit)

We report the correlations of PRO social support variables (MSPSS) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Social Support Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Social Support, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

In the health group with all participants, PRO *family* items Q8 (*talks about problems*) and Q11 (*willing to help make decisions*) correlated with the most TechROs.

In the group with healthy participants, PRO *friends* items, Q6 (*friends try to help*), Q9 (*friends share joys and sorrows*), and Q12 (*friends talk about problems*), had relatively more correlations with TechRos than PRO *significant other* or *family* items. Furthermore, the PRO *friends numeric score* had many correlations with TechROs.

In the group with diseased participants, PRO *family* Q4 (*family gives emotional help and support*) correlated negatively with TechRO absolute sedentary duration and Q12 (*friends talk about problems*) positively with the TechRO *steps* (Table 8, rows with Health: Diseased).

Table 8. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of social support (MSPSS scale) and TechROs (Fitbit wearable).

	PRO		TechRO			Correlation/Contour				
Health	**Source**	**Variable**	**Amount**	**Family**	**Variable**		**Lower**	r_S	**Higher**	
All	Significant other	Q2: shares joys and sorrows	Relative	CLR PA+S	Vigorous	+0.3	+0.7	+0.8		
All	Significant other	Q5: a real source of comfort	Relative	CLR PA+S	Vigorous	+0.4	+0.7	+0.8		
All	Significant other	Q10: cares about feelings	Relative	CLR PA+S	Vigorous	+0.5	+0.7	+0.8		
All	Family	Q3: tries to help	Relative	CLR PA+S	Fair		+0.3	+0.8	+0.7	
All	Family	Q8: talks about problems	Relative	CLR PA+S	Fair		+0.6	+0.8	+0.8	
All	Family	Q8: talks about problems	Relative	CLR PA+S	Vigorous	+0.6	+0.8	+0.8		
All	Family	Numeric sub-score	Relative	CLR PA+S	Fair		+0.3	+0.8	×	
Healthy	Significant other	Q1: around when in need	Absolute	Processed	Fair		×	−0.9	−0.6	
Healthy	Significant other	Q2: shares joys and sorrows	Absolute	Processed	Fair		×	−0.9	−0.7	−0.4
Healthy	Significant other	Q5: a real source of comfort	Absolute	Processed	Fair		×	−0.9	−0.6	
Healthy	Significant other	Q5: a real source of comfort	Relative	CLR PA+S	Fair		+0.4	+0.8	+0.6	
Healthy	Significant other	Q10: cares about feelings	Absolute	Processed	Fair		×	−0.8	−0.7	−0.7
Healthy	Significant other	Numeric sub-score	Absolute	Processed	Fair		×	−0.9	−0.6	−0.5
Healthy	Family	Q3: tries to help	Absolute	Processed	Fair		×	−0.8	−0.6	
Healthy	Family	Q3: tries to help	Relative	CLR PA+S	Fair	+0.5	+0.5	+0.9	+0.6	
Healthy	Family	Q8: talks about problems	Absolute	Processed	Fair		×	−0.8	−0.5	−0.4
Healthy	Family	Q8: talks about problems	Relative	CLR PA+S	Fair	+0.6	+0.5	+0.8	+0.6	
Healthy	Family	Q11: willing to help make decisions	Relative	CLR PA	Fair		+0.4	+0.8	×	
Healthy	Family	Numeric sub-score	Relative	CLR PA+S	Fair	+0.5	+0.4	+0.8	+0.4	
Healthy	Friends	Q9: share joys and sorrows	Absolute	Processed	Light		×	+0.8	+0.7	+ 0.4
Healthy	Friends	Q12: talk about problems	Absolute	Processed	Light		×	+0.8	+0.7	
Healthy	All	Categorical score	Absolute	Processed	Active			+0.8		
Healthy	All	Categorical score	Relative	CLR PA	Light		×	+0.8	×	
Healthy	All	Numeric score	Absolute	Processed	Light+Fair		+0.7	+0.8	×	
Healthy	All	Numeric score	Relative	CLR PA+S	Fair	+0.6	+0.5	+0.8	+0.4	
Diseased	Family	Q4: gives emotional help and support	Absolute	Processed	Sedentary			−0.8	×	
Diseased	Friends	Q12: talk about problems	Absolute	Raw	Steps			+0.8		

Color coding: from orange (weak correlation) to green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

Social Support Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (MSPSS) and TechRO variables (Fitbit) in Table 8.

In the health group with all participants, several PRO items related to the *significant other* social support, Q2 (*a special person shares joys and sorrows*), Q5 (*a special person is a real source of comfort*), and Q10 (*a special person cares about my feelings*) correlated strongly and with a broad contour with TechRO relative *vigorous* activity durations when including sleep (Table 8, rows with Health: All and PRO Source: Significant other). Also, several PRO *family* items, Q3 (*family tries to help*) and Q8 (*family talks about problems*) as well as the *family numeric sub-score* correlated strongly and with a broad contour with TechRO relative *fair* and *vigorous* activity durations when including sleep. These two strong co-calibrations only appeared as highlighted in the CLR PA+S family (Table 8, rows with Health: All and PRO Source: Family).

In the group with healthy participants, we observed numerous strong negative correlations with broad contours between numerous PRO items. Several are related to the *significant other* source: Q1 (*a special person is around when in need*), Q2 (*a special person shares joys and sorrows*), Q5 (*a special person is a real source of comfort*), and Q10 (*a special person cares about my feelings*) as well as the *significant other numeric sub-score* and the TechRO *fair* physical activity duration. However, we also observed a strong, positive correlation with a similarly sized contour with PRO item Q5 (*a special person is a real source of comfort*) and TechRO *fair* activity duration in the relative CLR PA+S family. These results indicate

that measuring daily sleep is necessary to co-calibrate this PRO source and TechRO physical activity intensity (Table 8, rows with Health: Healthy and PRO Source: Significant other).

Also, several PRO *family* items, Q3 (*family tries to help*), Q8 (*family talks about problems*), and Q11 (*family is willing to help make decisions*) correlated negatively with TechRO absolute *fair* activity, but positively with the relative duration at the same physical activity intensity (Table 8, rows with Health: Healthy and PRO Source: Family), yielding a similar interpretation.

Few PRO *friends* items such as Q9 (*friends share joys and sorrows*) and Q12 (*friends talk about problems*) correlated with broad contours with the TechRO absolute *light* physical activity duration (Table 8, rows with Health: Healthy and PRO Source: Friends).

Also, the PRO *categorical score* strongly correlated without contour with the TechRO absolute daily duration of physical activity (*active*) and the relative CLR PA *light* activity. The PRO *numeric score* also correlated with the TechRO absolute *light+fair* activity and relative CLR PA+S *fair* activity, indicating a positive relationship between social support and light to fair activity (Table 8, rows with Health: Healthy and PRO Source: All).

In the group with diseased participants, we only observed two isolated strong correlations. PRO *family* item Q4 (*gives emotional help and support*) correlated negatively with TechRO *sedentary* duration. PRO *friends* item Q12 (*talk about problems*) correlated positively with daily *steps* (Table 8, rows with Health: Diseased).

PRO items Q2, Q3, Q5, Q8, Q10, and the *numeric score* appeared in both groups of all and healthy participants. However, only Q8 maintained the correlation with TechRO *fair* physical activity across health groups. Q12 had strong correlations in both groups of healthy and diseased participants. However, the relationship was expressed through separate outcomes: *light* activity and *steps*, respectively (Table 8).

Social Support Outcomes Highlighted by Both Metrics

In the health group with all participants, PRO *friends* Q9 (*friends share joys and sorrows*) and Q12 (*friends talk about problems*) were highlighted as strongly correlated by both contour and total metrics, and thus informative for co-calibration with TechROs (Tables 6 and 8, rows with Health: All).

In the group with healthy participants, for the *significant other* and *family* sources of social support, Q10 (*a special person cares about my feelings*) and Q3 (*family tries to help*) appeared as informative with both metrics (Tables 6 and 8, rows with Health: Healthy).

Social Support Outcomes Interpretation

In the health group with all participants, several PRO items related to the *significant other* and *family* social support. They alternatively correlated with TechRO relative *fair* and *vigorous* activity: *family* items to the *fair* activity, and *significant other* items to the *vigorous* activity. All correlations resulted from relative TechROs including sleep. For this reason, the assessment of social support may benefit from the inclusion of sleep in the analysis.

In the group with healthy participants, the PRO social support from the *significant other* had negative correlations with TechRO *fair* activity in the absolute amount and positive correlations with *fair* activity in the relative amount (including sleep). This pattern was also pronounced for the items related to *family* social support. Sleep changed the ordering of durations throughout the day across the healthy participants. We argue for including sleep in the analysis of *significant other* and *family* social support for healthy seniors. Having *friends* who *share joys and sorrows* and, in general, *talk about problems*, associated with more light activity.

In the group with diseased participants, *emotional help* and *support* from the *family* associated with less *sedentary* time throughout the day. Also, having *friends* who *talk about problems* associated with more *steps*.

In general, the *significant other being a real source of comfort* appeared in most instances, followed by *having someone who cares about feelings*, then having someone who *shares joys and sorrows*, and then

(at a distance) having a special person *around when in need*. Having a *significant other* who is *a source of comfort* may serve as a proxy item for more frequent assessments of the relationships between *significant other* social support and physical activity at the *fair* to *vigorous* intensities.

Having a *family* that *tries to help*, *talks about problems*, and *wishes to help make decisions* appeared in three groups across metrics. However, *getting emotional help and support* from the family only appeared once. Frequent administrations of the MSPSS may choose to assess the relationships between *family* social support and *fair* physical activity by using only the first three items.

Having *friends* with whom to *talk about problems* appeared in three groups across metrics. Having *friends* who *try to help* and *share joys and sorrows* appeared less often with strong correlations and contours but had numerous correlations in total. We argue that *counting on friends when things go wrong* is a less prominent item in assessing relationships between *friends* social support and physical activity.

3.4.3. coQoL for Anxiety and Depression (GADS vs. Fitbit)

We report the correlations of PRO anxiety and depression (GADS) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Anxiety and Depression Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Anxiety and depression, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

In the health group with all participants, PRO *anxiety* item Q8A (*worried about own health*), as well as PRO *depression* items Q1D (*lacking energy*) and Q6D (*lost weight due to poor appetite*), recorded the most correlations with TechROs (Table 6, row with Outcome: Anxiety and depression, Health: All).

In the group with healthy participants, PRO item Q2D (*lost interest in things*) had the most correlations (Table 6, row with Outcome: Anxiety and depression, Health: Healthy).

In the group with diseased participants, PRO item Q2A (*worrying a lot*) had the most correlations with TechROs (Table 6, row with Outcome: Anxiety and depression, Health: Diseased).

Anxiety and Depression Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (GADS) and TechRO variables (Fitbit) in Table 9.

In the health group with all participants, PRO *anxiety* item Q5A (*sleeping poorly*) correlated strongly with a broad contour with TechRO relative CLR PA+S *light* physical activity. We found other isolated correlations for *anxiety*. PRO item Q3A (*irritable*) correlated with the TechRO relative *vigorous* activity. PRO item Q7A (*trembling [. . .]*) negatively correlated with the TechRO daily *active* duration. PRO *depression* items Q1D (*lacking energy*) and Q6D (*lost weight due to poor appetite*) had isolated correlations. The PRO *numeric score* had a strong correlation with the TechRO relative *sleep* duration (Table 9, rows with Health: All).

In the group with healthy participants, PRO *anxiety* item Q7A (*trembling [. . .]*) correlated positively with TechRO *vigorous* activity and negatively with TechRO *light* and *light+fair* activity durations (the last with a broad contour) in both absolute and relative families. PRO item Q7A correlated negatively with the total daily *active* duration. PRO item Q3A (*irritable*) correlated negatively with total daily *active* duration. PRO *depression* items Q2D (*lost interest in things*) and Q9D (*worse in the morning*) had isolated correlations, the first negative with TechRO relative CLR PA *light* activity duration, and the second with TechRO relative CLR PA+S *sedentary* duration. PRO item Q6D (*lost weight due to poor appetite*) recorded a positive correlation as well, with TechRO relative *sleep* duration (Table 9, rows with Health: Healthy).

In the group with diseased participants, we did not observe strong correlations ($r_S \geq 0.8$) by using the contour metric (Table 9, rows with Health: Diseased).

PRO items Q3A, Q7A, and Q6D appeared in both groups with all and healthy participants. However, only Q7A kept the same strong correlation against total daily *active* duration in the two groups (Table 9).

Table 9. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of anxiety and depression (GADS scale) and TechROs (Fitbit wearable).

	PRO		TechRO			Correlation/Contour			
Health	**Outcome**	**Variable**	**Amount**	**Family**	**Variable**	**Lower**	r_S	**Higher**	
All	Anxiety	Q3A: irritable	Relative	CLR PA	Vigorous	×	+0.8		
All	Anxiety	Q5A: sleeping poorly	Relative	CLR PA+S	Light	+0.5	+0.8	+0.5	+0.3
All	Anxiety	Q7A: trembling	Absolute	Processed	Active		−0.8		
All	Depression	Q1D: lacking energy	Relative	CLR PA+S	Vigorous	×	−0.8		
All	Depression	Q6D: lost weight due to poor appetite	Relative	CLR PA+S	Light	×	+0.8	×	
All	Both	Numeric score	Relative	CLR PA+S	Sleep		+0.8		
Healthy	Anxiety	Q3A: irritable	Absolute	Processed	Active		−0.8		
Healthy	Anxiety	Q7A: trembling	Absolute	Processed	Light+fair	−0.5	−0.8	−0.5	
Healthy	Anxiety	Q7A: trembling	Absolute	Processed	Vigorous	×	+0.8		
Healthy	Anxiety	Q7A: trembling	Absolute	Processed	Active		−0.8		
Healthy	Anxiety	Q7A: trembling	Relative	CLR PA	Light	×	−0.8	×	
Healthy	Anxiety	Q7A: trembling	Relative	CLR PA+S	Vigorous	×	+0.8		
Healthy	Depression	Q2D: lost interest in things	Relative	CLR PA	Light	×	−0.8	×	
Healthy	Depression	Q6D: lost weight due to poor appetite	Relative	CLR PA+S	Sleep		+0.8		
Healthy	Depression	Q9D: worse in the morning	Relative	CLR PA+S	Sedentary		+0.8	×	

Color coding: from orange (weak correlation) to green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

Anxiety and Depression Outcomes Highlighted by Both Metrics

In the health group with all participants, PRO items Q1D (*lacking energy*) and Q6D (*lost weight due to poor appetite*) were highlighted by both metrics (Tables 6 and 9, rows with Health: All).

For healthy participants, PRO item Q2D (*lost interest in things*) appeared in both metrics as well (Tables 6 and 9, rows with Health: Healthy).

Anxiety and Depression Outcomes Interpretation

In the health groups with all and healthy participants, *irritability* and *trembling* may expediently assess *anxiety* while having *lost interest in things* and *losing weight due to poor appetite* may assess *depression*. Follow-up investigations may establish whether the health state is momentary or deteriorating over time.

PRO *Trembling, tingling, dizziness, sweating, diarrhoea, or passing urine* yielded numerous correlations for healthy participants: negative correlations with TechRO *light*, *light+fair*, and total daily *active* duration as well as a positive correlation with *vigorous* physical activity duration. When a daily life monitor observed a gradual replacement of *light* to *fair* activity with *vigorous* activity (as reported by the wearable), it may be worth investigating whether an otherwise healthy participant also becomes gradually more anxious (by using items).

In the group with healthy participants, a decrease in *light* physical activity may indicate that the participants experience an increase in *depression*. Researchers can then assess this hypothesis by administering, e.g., the corresponding item in the EQ-5D-3L scale. A similar process could be employed for all seniors by longitudinally monitoring the *sleep* duration relative to the 24 h of the day, based on the corresponding strong correlations between the *numeric score* and the relative *sleep* duration. In the case of increasingly longer *sleep*, the participant may enter a state of *anxiety* or *depression*.

In general, *depression* and *anxiety* positively associated with the *sedentary* duration, in both absolute and relative TechRO families, especially for participants who self-report disease. The two items in the

scale referring to sleep may provide additional insights towards not only the anxiety and depression status of the participant, but also sleep quality.

3.4.4. coQoL for Mediterranean Nutrition (PREDIMED vs. Fitbit)

We report the correlations of PRO Mediterranean nutrition variables (PREDIMED) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Mediterranean Nutrition Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Mediterranean nutrition, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

In the health group with all participants, the PRO *categorical score*, *numeric score* and items Q12 (*nuts use*) and Q14 (*sofrito use*) had the most correlations with TechROs (Table 6, rows with Outcome: Mediterranean nutrition, Health: All).

In the group with healthy participants, PRO item Q4 (*fruit use*) and the *categorical score* had the most correlations with TechROs (Table 6, rows with Outcome: Mediterranean nutrition, Health: Healthy).

In the group with diseased participants, we only observed PROs with reduced numbers of correlations with TechROs across families (Table 6, rows with Outcome: Mediterranean nutrition, Health: Diseased).

The *categorical score* is the only PRO that appeared with numerous correlations in the two groups with all and healthy participants (Table 6).

Mediterranean Nutrition Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (PREDIMED) and TechRO variables (Fitbit) in Table 10.

Table 10. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of Mediterranean nutrition (PREDIMED scale) and TechROs (Fitbit wearable).

	PRO	TechRO			Correlation/Contour		
Health	**Variable**	**Amount**	**Family**	**Variable**	**Lower**	r_S	**Higher**
All	Q12: nuts use	Absolute	Processed	Fair	×	−0.9	×
All	Q12: nuts use	Relative	CLR PA+S	Light	+0.6	+0.8	×
All	Numeric score	Absolute	Processed	Vigorous	−0.7	−0.8	
All	Numeric score	Relative	CLR PA+S	Light	+0.6	+0.8	+0.6
Healthy	Q3: vegetables use	Relative	CLR PA	Fair	×	−0.8	×
Healthy	Q3: vegetables use	Relative	CLR PA+S	Fair	×	−0.8	−0.4
Diseased	Q5: red meat, hamburger, or meat use	Absolute	Raw	Energy		+0.8	
Diseased	Q11: commercial sweets or pastries use	Absolute	Raw	Heart rate		+0.8	

Color coding: from orange (weak correlation) to green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

In the health group with all participants, PRO item Q12 (*nuts use*) had an isolated negative correlation with the TechRO absolute *fair* activity, but a positive correlation (with a contour) with the TechRO relative CLR PA+S *light* activity. The PRO *numeric score* also registered two correlations with contours: negative with TechRO absolute *vigorous* activity duration and positive with TechRO relative CLR PA+S *light* activity duration (Table 10, rows with Health: All).

In the group with healthy participants, PRO item Q3 (*vegetables use*) correlated negatively with the TechRO relative *fair* activity in both CLR PA and CLR PA+S families (Table 10, rows with Health: Healthy). While the two correlations had no contour, their presence in both families highlights an effect.

In the group with diseased participants, PRO item Q5 (*red meat, hamburger, or meat use*) correlated positively with TechRO *energy* expenditure. For the same group, PRO item Q11 (*commercial sweets or pastries use*) correlated positively with TechRO *heart rate* (Table 10, rows with Health: Diseased).

Mediterranean Nutrition Outcomes Highlighted by Both Metrics

For all participants, PRO item Q12 (*nuts use*) and the *numeric score* were highlighted by both metrics (Tables 6 and 10, rows with Health: All).

Mediterranean Nutrition Outcomes Interpretation

In the health group with all participants, the nutrition *numeric score* associated with the relative *sleep* duration, and *using nuts* had a similar correlation (both correlations with contours). Further studies may assess whether this item can be administered independently of the full scale (for the *numeric score*) to assess the relationship between (mal)nutrition and *light* physical activity in seniors.

With regards to poor nutrition choices and their potentially magnified effects on people with mild disease, the *consumption of red meat and hamburgers* by participants with mild disease correlated with higher *energy* expenditure. The consumption of *commercial sweets or pastries* also associated with an increased *heart rate*.

The PRO *numeric* and *categorical scores* correlated with numerous TechROs, indicating a replacement of *fair* to *vigorous* activity with the *light* activity.

Participants from Spain had on average more adherence than those from Hungary (Appendix C.1.1), making the country of residence a potential confounder for the relationships above.

3.4.5. coQoL for Nutrition (SelfMNA vs. Fitbit)

We report the correlations of PRO nutrition variables (SelfMNA) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Nutrition Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Nutrition, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

For all health groups, we found PROs correlated with few TechROs when compared to other outcomes (Table 6, row with Outcome: Nutrition, Health: All).

In the groups with all participants and the healthy, the PRO *categorical score* had the most correlations (Table 6, row with Outcome: Nutrition, Health: Healthy).

In the group with diseased participants, PRO items Q1 (*food intake declined*) and Q2 (*weight lost*) recorded the most correlations with TechROs (Table 6, row with Outcome: Nutrition, Health: Diseased).

The *categorical score* is the only PRO that appeared in two health groups: the group with all participants and the group with healthy participants (Table 6).

Nutrition Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (SelfMNA) and TechRO variables (Fitbit) in Table 11.

We only found strong correlations ($r_S \geq 0.8$) in the group with diseased participants. PRO items Q1 (*food intake declined*) and Q2 (*weight lost*) correlated negatively with the TechRO relative *sleep* duration. PRO item Q4 (*stressed or severely ill*) correlated negatively with the TechRO absolute *sedentary* duration (Table 11).

Table 11. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of nutrition (SelfMNA scale) and TechROs (Fitbit wearable).

	PRO	TechRO			Correlation/Contour		
Health	Variable	Amount	Family	Variable	Lower	r_S	Higher
Diseased	Q1: food intake declined	Relative	CLR PA+S	Sleep		-0.8	
Diseased	Q2: weight lost	Relative	CLR PA+S	Sleep		-0.8	
Diseased	Q4: stressed or severely ill	Absolute	Processed	Sedentary		-0.8	×

Green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

Nutrition Outcomes Highlighted by Both Metrics

In the group with diseased participants, PRO items Q1 (*food intake declined*) and Q2 (*weight lost*) were highlighted by both metrics (Tables 6 and 11, rows with Health: Diseased).

Nutrition Outcomes Interpretation

In the health group with all participants, the PRO *categorical score* correlated with numerous TechROs. In general, better nutrition coincided with less *sedentary* and light *physical* activity and more *fair* and *vigorous* physical activity. In the group with healthy participants, both *numeric* and *categorical scores* exhibited this pattern (Appendix C.2).

In the group with diseased participants, a long-term decrease in *sleep* duration may indicate a *decline in food intake* or a *loss of weight*—two outcomes that appeared in both metrics and may lead to malnutrition.

3.4.6. coQoL for Memory (MFE vs. Fitbit)

We report the correlations of PRO memory variables (MFE) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Memory Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Memory, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

In the health group with all participants, the PRO items that correlated with the most TechROs were Q12 (*having difficulty picking up a new skill*), Q14 (*forgetting to do planned things*), and Q6 (*forgetting the time of events*) (Table 6, rows with Outcome: Memory and Health: All).

In the group with healthy participants, PRO items Q6 (*forgetting the time of events*), Q15 (*forgetting details of done things*), Q12 (*having difficulty picking up a new skill*), and Q14 (*forgetting to do planned things*) correlated with the most TechROs (Table 6, rows with Outcome: Memory and Health: Healthy).

In the group with diseased participants, PRO items Q13 (*having a word on the tip of the tongue*) and Q25 (*getting lost in often visited place*) had the most correlations (Table 6, rows with Outcome: Memory and Health: Diseased).

PRO items Q12 (*having difficulty picking up a new skill*) and Q14 (*forgetting to do planned things*) were the only outcomes that had numerous correlations with TechROs across two groups: all and healthy (Table 6).

Memory Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (MFE) and TechRO variables (Fitbit) in Table 12.

In the health group with all participants, there was only one strong correlation with contour between PRO item Q24 (*forgetting where things are normally kept*) and PRO *fair* activity in the CLR PA family. The PRO *numeric score* had a negative correlation with the TechRO total daily *active* duration. PRO item Q7 (*completely forgetting to take things*) had a strong correlation with TechRO

relative *sleep* duration. PRO items Q12 (*having difficulty picking up a new skill*) and Q13 (*finding a word on the tip of the tongue*) had negative and positive relations with TechRO relative *light* and *fair* CLR PA+S activity durations, respectively (Table 12, rows with Health: All).

Table 12. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of memory (MFE scale) and TechROs (Fitbit wearable).

	PRO	TechRO			Correlation/Contour		
Health	**Variable**	**Amount**	**Family**	**Variable**	**Lower**	r_S	**Higher**
All	Q7: completely forgetting to take things	Relative	CLR PA+S	Sleep		+0.8	
All	Q12: having difficulty picking up a new skill	Relative	CLR PA+S	Light	×	−0.8	×
All	Q13: finding a word on the tip of the tongue	Relative	CLR PA+S	Sleep		+0.8	
All	Q24: forgetting where things are normally kept	Relative	CLR PA	Fair	×	+0.8	×
All	Q24: forgetting where things are normally kept	Relative	CLR PA+S	Fair	×	−0.8	−0.3
All	Numeric score	Absolute	Processed	Active		−0.8	
Healthy	Q7: completely forgetting to take things	Relative	CLR PA+S	Sleep		+0.8	
Healthy	Q10: letting ramble about unimportant things	Absolute	Processed	Light+fair	×	−0.8	×
Healthy	Q14: forgetting to do planned things	Absolute	Processed	Fair+vigorous	×	+0.8	+0.8
Healthy	Q14: forgetting to do planned things	Absolute	Processed	Vigorous	+0.8	+0.8	
Healthy	Q16: forgetting the topic of an ongoing conversation	Absolute	Processed	Fair	×	−0.8	−0.4
Healthy	Q24: forgetting where things are normally kept	Relative	CLR PA+S	Fair	×	−0.8	×
Healthy	Numeric score	Relative	CLR PA	Fair	×	−0.8	×
Diseased	Q1: forgetting objects put	Relative	CLR PA+S	Vigorous	−0.7	−0.8	
Diseased	Q6: forgetting the time of events	Absolute	Raw	Heart rate		+0.8	
Diseased	Q6: forgetting the time of events	Absolute	Processed	Light	+0.7	+0.8	×
Diseased	Q6: forgetting the time of events	Absolute	Processed	Sleep		−0.8	
Diseased	Q8: being reminded about things	Absolute	Processed	Light+fair	+0.6	+0.8	×
Diseased	Q9: reading anew something already read	Absolute	Processed	Sleep		−0.8	
Diseased	Q13: finding a word on the tip of the tongue	Absolute	Processed	Active		−0.8	
Diseased	Q13: finding a word on the tip of the tongue	Relative	CLR PA+S	Sedentary		+0.8	+0.7
Diseased	Q18: forgetting to tell somebody something important	Absolute	Processed	Fair	×	−0.8	−0.8 −0.8
Diseased	Q18: forgetting to tell somebody something important	Absolute	Processed	Fair+vigorous	−0.8	−0.8	−0.8
Diseased	Q18: forgetting to tell somebody something important	Absolute	Processed	Vigorous	−0.8	−0.8	−0.8
Diseased	Numeric score	Absolute	Processed	Active		−0.8	

Color coding: from orange (weak correlation) to green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

In the group with healthy participants, PRO item Q14 (*forgetting to do planned things*) had a contour of two strong correlations with TechRO *fair+vigorous* and *vigorous* activity. PRO item Q16 (*forgetting the topic of an ongoing conversation*) had a strong correlation with contour TechRO absolute *fair* activity duration. PRO items Q10 (*letting ramble about unimportant things*) and Q24 (*forgetting where things are normally kept*) had isolated negative correlations with TechRO *fair* activity duration. PRO item Q7 (*completely forgetting to take things*) recurred in correlating strongly with *sleep*. The *numeric score* also correlated negatively with TechRO relative CLR PA *fair* activity duration (Table 12, rows with Health: Healthy).

In the group with diseased participants, PRO item Q18 (*forgetting to tell somebody something important*) had a broad contour with the TechRO *fair*, *fair+vigorous*, and *vigorous* physical activity duration. PRO item Q6 (*forgetting the time of events*) had a positive correlation with the TechRO *heart rate*, a positive correlation (having a contour) with the *light* activity, and a negative correlation with the *sleep* duration. PRO item Q1 (*forgetting objects put*) had a negative correlation (contour) with the TechRO relative *vigorous* activity in the PA+S family. Q13 (*finding a word on the tip of the tongue*) correlated negatively with TechRO daily *active* duration and positively with relative *sedentary* duration in the CLR PA+S family. Q8 (*being reminded about things*) had a positive correlation with the TechRO *light+fair* activity duration. The PRO *numeric score* correlated negatively with the TechRO total *active* duration (Table 12, rows with Health: Diseased).

PRO items Q7 (*completely forgetting to take things*) and Q24 (*forgetting where things are normally kept*), as well as the *numeric score*, appeared in both groups with all and healthy participants. Items Q7 and Q24 maintained the strong correlations between groups: positive with *sleep* duration and negative with relative *fair* activity. The *numeric score* expressed the inverse relation with physical activity in different ways depending on the health status. For all participants and the mildly diseased, it had a negative correlation with the total daily *active* duration. For the healthy participants, it had a negative correlation with the relative *fair* activity duration (Table 12).

Memory Outcomes Highlighted by Both Metrics

In the health group with all participants, Q12 (*having difficulty picking up a new skill*) was highlighted by both metrics as an informative PRO for memory (Tables 6 and 12, rows with Health: All).

In the group with healthy participants, PRO item Q14 (*forgetting to do planned things*) was informative in both metrics (Tables 6 and 12, rows with Health: Healthy).

In the group with diseased participants, PRO item Q13 (*finding a word on the tip of the tongue*) was informative through both metrics (Tables 6 and 12, rows with Health: Diseased).

Memory Outcomes Interpretation

In the health group with all participants, the memory *numeric score* strongly associated with shorter durations of any physical activity during the day. A negative correlation with relative *fair* physical activity also reflected this pattern in the group with healthy participants. A decrease in *active* duration may provide an opportunity for a long-term monitoring system to assess whether an otherwise healthy senior is experiencing a gradual increase in memory failures.

In the groups with all participants and the healthy, *forgetting where things are normally kept* associated positively with *fair* physical activity; however, only when accounting for sleep as well.

In the group with diseased participants, *forgetting to tell somebody something important* associated with numerous TechROs, suggesting a replacement of *fair* and *vigorous* activity durations with *sedentary* and *light* duration throughout the day. By observing this TechRO pattern longitudinally in time, a study may administer this item towards assessing memory failures. *Finding a word is on the tip of the tongue* is another PRO item that also correlated with TechRO *sedentary* duration and negatively correlated with daily *active* duration. Further research may investigate the reliability of a more frequent assessment than the MFE scale consisting of the items above for seniors with mild disease.

3.4.7. coQoL for Sleep (PSQI vs. Fitbit)

We report the correlations of PRO sleep variables (PSQI) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Sleep Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Sleep, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

In the health group with all participants, PRO items Q7 (*trouble staying awake driving, eating, socialising*) and Q4 (*duration of actual sleep*), followed by the *daily dysfunction numeric sub-score*, had the most correlations with TechROs across families (Table 6, rows with Outcome: Sleep and Health: All).

In the group with healthy participants, PRO items Q4 (*duration of actual sleep*), Q5C (*trouble sleeping due to using the bathroom*), Q7 (*trouble staying awake driving, eating, socialising*) had the most correlations with TechROs, followed by the *daily dysfunction numeric sub-score* (Table 6, rows with Outcome: Sleep and Health: Healthy).

In the group with diseased participants, the PROs that correlated with the most TechROs had relatively fewer correlations. The *daily dysfunction numeric sub-score* and Q6 (*duration of actual sleep*) registered the most correlations (Table 6, rows with Outcome: Sleep and Health: Diseased).

The PRO *daily dysfunction numeric sub-score* had numerous correlations in all three health groups. The PRO item Q4 (*duration of actual sleep*) appeared in the groups with all participants and the healthy (Table 6).

Sleep Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (PSQI) and TechRO variables (Fitbit) in Table 13.

Table 13. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of sleep (PSQI scale) and TechROs (Fitbit wearable).

	PRO	TechRO			Correlation/Contour				
Health	**Variable**	**Amount**	**Family**	**Variable**	**Lower**		r_S	**Higher**	
All	Q5A: trouble sleeping due to not getting to sleep	Relative	CLR PA+S	Sleep			+0.8		
All	Q5E: trouble sleeping due to coughing or snoring loudly	Relative	CLR PA	Vigorous		−0.5	−0.8		
All	Q5F: trouble sleeping due to feeling too cold	Relative	CLR PA+S	Light		+0.6	+0.8	+0.6	
All	Q7: trouble staying awake while driving, eating, socializing	Relative	CLR PA	Light		−0.5	−0.8	×	
All	Q7: trouble staying awake while driving, eating, socializing	Relative	CLR PA+S	Sleep			−0.8		
All	Latency numeric sub-score	Relative	CLR PA+S	Sleep			+0.8		
All	Efficiency numeric sub-score	Relative	CLR PA	Fair		×	+0.8	×	
All	Daily dysfunction numeric sub-score	Absolute	Processed	Vigorous	+0.5	+0.5	+0.8		
All	Daily dysfunction numeric sub-score	Relative	CLR PA	Light		−0.6	−0.8	×	
All	Daily dysfunction numeric sub-score	Relative	CLR PA+S	Sleep			−0.8		
Healthy	Q2: duration taken to fall asleep	Relative	CLR PA+S	Sleep			+0.8		
Healthy	Q3: time gotten up in the morning	Absolute	Raw	Energy			−0.8		
Healthy	Q5A: trouble sleeping due to not getting to sleep	Relative	CLR PA+S	Sleep			+0.8		
Healthy	Q5B: trouble sleeping due to waking up in the middle of the night	Relative	CLR PA+S	Vigorous		×	+0.8		
Healthy	Q5C: trouble sleeping due to using the bathroom	Absolute	Processed	Light+Fair		−0.5	−0.8	×	
Healthy	Q5C: trouble sleeping due to using the bathroom	Relative	CLR PA	Light		×	−0.8	−0.5	−0.6
Healthy	Q5E: trouble sleeping due to coughing or snoring loudly	Relative	CLR PA+S	Light		×	−0.8	×	
Healthy	Q11: duration stayed in bed	Relative	CLR PA+S	Sleep			+0.8		
Healthy	Numeric score	Absolute	Processed	Fair+vigorous		×	+0.8	+0.6	
Healthy	Latency numeric sub-score	Relative	CLR PA+S	Sleep			+0.8		
Healthy	Efficiency numeric sub-score	Relative	CLR PA	Fair		×	+0.8	×	
Diseased	Q1: time gone to bed at night	Absolute	Processed	Sleep			−0.8		
Diseased	Q4: duration of actual sleep	Absolute	Processed	Fair		×	+0.8	+0.8	+0.9
Diseased	Q4: duration of actual sleep	Absolute	Processed	Fair+vigorous		+0.8	+0.8	+0.9	
Diseased	Q4: duration of actual sleep	Absolute	Processed	Vigorous	+0.8	+0.8	+0.9		
Diseased	Q5B: trouble sleeping due to waking up in the middle of the night	Absolute	Raw	Energy			−0.8		
Diseased	Q5C: trouble sleeping due to using the bathroom	Absolute	Raw	Energy			−0.8		

Color coding: from orange (weak correlation) to green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

In the health group with all participants, PRO *sleep disturbance* item Q5A (*trouble sleeping due to not getting to sleep*) correlated positively with TechRO relative *sleep* duration. PRO items Q5E (*trouble sleeping due to coughing or snoring loudly*) and Q5F (*trouble sleeping due to feeling too cold*) correlated with TechRO relative *vigorous* activity duration (negative, CLR PA family) and *light* activity duration (positive, CLR PA+S family), respectively. PRO item Q7 (*trouble staying awake while driving, eating, socialising*) correlated negatively with TechRO relative *sleep* duration and *light* activity durations. Two *numeric sub-scores* yielded correlations with relative *sleep*: *latency* (positive) and *daily dysfunction* (negative). The *daily dysfunction numeric sub-score* also correlated with TechRO *vigorous* activity (broad contour) and the relative *light* activity (contour). The *efficiency numeric sub-score* had an isolated correlation with TechRO *fair* activity (Table 13, rows with Health: All).

In the group with healthy participants, numerous PROs correlated with TechRO *sleep*: Q2 (*duration to fall asleep*), Q5A (*trouble sleeping due to not getting to sleep*), Q11 (*duration stayed in bed*), and the *latency numeric sub-score*. Among the *sleep disturbance* items, Q5C (*trouble sleeping due to using the bathroom*) had two contoured correlations: negative with *light+fair* and *light* activity (the latter with a broad contour) in absolute and relative CLR PA families, respectively. The PRO *efficiency numeric sub-score* correlated

again with TechRO *fair* activity. The *numeric score* correlated positively (and having a contour) with *fair+vigorous* activity (Table 13, rows with Health: Healthy).

In the group with diseased participants, PRO item Q4 (*duration of actual sleep*) registered a broad contour of 3 strong correlations (including r_S = 0.9) with *fair*, *fair+vigorous*, and *vigorous* TechRO absolute durations. PRO item Q1 (*time gone to bed at night*) correlated inversely with the TechRO absolute *sleep* duration. *Sleep disturbance* items Q5B (*trouble sleeping due to waking up in the middle of the night*) and Q5C (*trouble sleeping due to using the bathroom*) correlated negatively with *energy* expenditure (Table 13, rows with Health: Diseased).

PRO items Q5A (*trouble sleeping due to not getting to sleep*) and Q5E (*trouble sleeping due to coughing or snoring loudly*), and the *latency* and *efficiency numeric sub-scores* appeared for the groups with all participants and the healthy. Q5A and the *latency numeric sub-score* maintained a strong correlation with the TechRO *sleep* duration. The *efficiency numeric sub-score* maintained the strong correlation with the *fair* activity. Q5E had an inverse relation with TechRO physical activity across these two groups, but expressed through negative correlations with the relative *vigorous* duration and the relative *light* duration, respectively. Q5C (*trouble sleeping due to using the bathroom*) was highlighted in both healthy and diseased groups, but expressed an inverse relation with physical activity through different outcomes: *light-fair* activity duration and *energy* expenditure, respectively (Table 13).

Sleep Outcomes Highlighted by Both Metrics

In the health group with all participants, PRO item Q7 (*trouble staying awake driving, eating, socialising*) appeared as informative in both metrics (Tables 6 and 13, rows with Health: All).

In the group with healthy participants, Q5C (*trouble sleeping due to using the bathroom*) was an informative PRO item that appeared in both metrics (Tables 6 and 13, rows with Health: Healthy).

Sleep Outcomes Interpretation

Several PRO items strongly correlated with sleep-specific TechROs. In the health group with all participants, *having trouble sleeping due to not being able to get to sleep* as well as the *sleep latency numeric sub-score* correlated with relative *sleep* duration while *having trouble staying awake while driving, eating, or socialising* as well as the *daily dysfunction numeric sub-score* correlated negatively with relative *sleep* duration. In the group with healthy participants, the *duration to fall asleep*, *having trouble sleeping due to not getting to sleep*, the *duration to stay in bed*, and the *latency numeric sub-score* correlated with longer relative *sleep* during the day. In the group with diseased participants, only the *time gone to bed at night* correlated negatively with absolute *sleep* duration. Studies assessing sleep in healthy adults may benefit from the monitoring of the entire day, not only the sleep duration, to find a higher amount of significant outcomes.

In the health group with all participants, PRO decreased sleep quality correlated negatively with TechRO relative *light* and *vigorous* activity. In the group with healthy participants, the *sleep efficiency numeric sub-score* correlated with the relative *fair* activity, and *using the bathroom* correlated negatively with relative *light* physical activity (with a broad contour). In the group with diseased participants, the *duration of actual sleep* correlated with absolute *fair*, *fair+vigorous*, and *vigorous* durations. *Having trouble sleeping due to waking up in the middle of the night* may be an indicator of already low sleep quality in participants with mild disease.

3.4.8. coQoL for Health-Related Quality of Life (EQ-5D-3L vs. Fitbit)

We report the correlations of PRO health-related Quality of Life variables (EQ-5D-3L) with TechRO variables (Fitbit) by using the *total* and *contour* metrics.

Health-Related Quality of Life Outcomes by Total Numbers of Correlations

Table 6, rows with Outcome: Quality of Life, enumerates the PROs that correlated with the most TechROs ($r_S \geq 0.5$) across all families by health group.

In the health group with all participants, the PRO items with the most correlations were the *health score* and Q4 (*pain/discomfort*). The items in this scale had relatively fewer correlations than the other scales such as social support (MSPSS) or memory (MFE) (Table 6, rows with Outcome: Quality of Life and Health: All).

In the group with healthy participants, PRO item Q4 (*pain/discomfort*) had the most correlations with TechROs (Table 6, row with Outcome: Quality of Life and Health: Healthy).

In the group with diseased participants, PRO item Q5 (*anxiety/depression*) had the most correlations with TechROs (Table 6, row with Outcome: Quality of Life and Health: Diseased).

Q4 (*pain/discomfort*) was the only PRO item that appeared in two groups: the group with all participants and the group with the healthy (Table 6).

Health-Related Quality of Life Outcomes by Contours of Correlations

We report the strong correlations ($r_S \geq 0.8$) and their contours between PRO variables (EQ-5D-3L) and TechRO variables (Fitbit) in Table 14.

Table 14. Summary of found strong and significant Spearman rank correlations ($r_S \geq 0.8$) between PROs of health-related Quality of Life (EQ-5D-3L scale) and TechROs (Fitbit wearable).

	PRO		TechRO			Correlation/Contour		
Health	**Domain**	**Variable**	**Amount**	**Family**	**Variable**	**Lower**	r_S	**Higher**
Diseased	Anxiety/depression	Q5: anxiety/depression	Absolute	Processed	Sedentary		+0.8	×

Color coding: green (strong correlation). × depicts an absent significant correlation of the same sign next to the strong correlation.

We only found one strong correlation in the group of participants with mild disease, between the PRO *depression and anxiety* item (Q5) and the TechRO absolute *sedentary* duration (Table 14).

Health-Related Quality of Life Outcomes Highlighted by Both Metrics

In the group with diseased participants, Q5 (*anxiety/depression*) recurred in both metrics (Tables 6 and 14, rows with Health: Diseased).

Health-Related Quality of Life Outcomes Interpretation

The PRO *health state today* correlated with numerous TechROs, in particular with a replacement of *vigorous* physical activity duration with *sleep*, *sedentary*, and *fair* durations across all participants, with a replacement of *fair* and *vigorous* durations with *light* activity for the healthy, and with a decrease in *fair* and *vigorous* activity among the diseased (Appendix C.2).

Pain and discomfort also had numerous correlations with TechROs, but only for the groups with all participants and the healthy. In participants with mild disease, having *anxiety/depression* correlated with *sedentary* physical activity. An increase in sedentary duration for participants with already existing mild disease may be an indication of decreased quality of life on the *anxiety/depression* domains which, in the affirmative, could be further assessed by administering specialized scales.

3.5. Use Case Examples for coQoL

The coQoL method allows for the in-depth analysis of the results both in terms of measured outcomes and individual participants. We provide two examples below, pertaining to longitudinal data (Section 3.5.1) and the story of a participant (Section 3.5.2).

3.5.1. Longitudinal Data Example

We exemplify a very strong correlation (r_S = 0.9) between PROs and TechROs, to report how the interval and leeway durations influenced the correlations. In healthy participants, the MSPSS item Q3 (*family is trying to help*, PRO) correlated the strongest with the Fitbit fair physical activity

duration in the CLR PA+S family, TechRO) for the TechRO aggregation interval of 28 days with a decreasing pattern as the leeway increases. Table 15 presents the resulting gradients of correlations for all combinations of TechRO aggregation interval-leeway durations and the TechRO raw data that yielded the strongest correlation. Table 16 depicts the raw results. In this table, the relative *fair* column is a centred log-ratio that has both negative (for less relative *fair* activity) and positive quantities (for more relative *fair* activity).

Table 15. Gradient of correlations by interval durations (columns) and leeways (rows) in days.

	7	14	21	28	60	90	120
0	-0.09^{11}	0.19^{9}	0.13^{9}	0.88^{9}	0.44^{8}	-0.25^{8}	0.14^{8}
7	-0.27^{16}	0.19^{9}	0.13^{9}	0.88^{9}	0.44^{8}	-0.36^{9}	0.14^{8}
14	-0.07^{23}	0.16^{16}	-0.03^{10}	0.88^{9}	0.44^{8}	-0.36^{9}	0.14^{8}
21	-0.07^{23}	-0.08^{20}	-0.14^{16}	0.92^{11}	0.44^{8}	-0.36^{9}	0.27^{9}
28	-0.07^{23}	-0.08^{20}	0.01^{17}	0.61^{13}	0.19^{9}	-0.36^{9}	0.27^{9}
60	-0.07^{23}	-0.09^{23}	-0.13^{21}	0.57^{20}	0.17^{10}	-0.36^{9}	0.27^{9}
90	-0.09^{24}	-0.06^{24}	-0.16^{22}	0.48^{21}	-0.08^{14}	-0.13^{10}	0.27^{9}
120	-0.06^{25}	-0.06^{24}	-0.16^{22}	0.48^{21}	-0.14^{15}	0.09^{16}	-0.10^{12}

Color coding: from yellow (weaker correlations) to green (stronger correlations). Superscript depicts sample size. Subscript depicts sign. All correlations are shown. Only significant correlations are highlighted.

Table 16. Raw data for a 28-day interval and a 21-day leeway that yielded the highest correlation (0.92).

Participant ID	Wave	Q3 (PRO)	Fair (TechRO)
617	2	4	−1.49
419	1	5	−1.54
419	2	5	−1.48
643	2	6	−1.24
793	3	6	+1.05
170	3	6	+1.49
569	1	7	+2.10
133	2	7	+1.73
569	2	7	+2.09
133	3	7	+1.69
569	3	7	+1.88

Color coding: from orange (lower values) to yellow to green (higher values).

3.5.2. Participant Story Example

Participant 169 is a 69-year-old female from Hungary who self-reported mild disease. She has a university degree, lives with her partner (no children), does not smoke, and drinks alcohol daily. She is a diligent responder who answered in all three waves of our study, wore the Fitbit for 794 days from which 141 were valid.

When aligning the numeric scores from the PRO scales and the TechROs (Table 17), Wave 1 (mid-2018) had the worst PRO depression and anxiety, (close to the worst) memory, and sleep as well as (close to) the worst TechRO sedentary duration, light activity duration, (close) fair activity, and vigorous activity duration. Wave 2 (end-2018 and start-2019) had the least adequate PRO physical activity, adherence to the Mediterranean diet, memory, sleep, and quality of life, reflected in the least adequate TechRO energy expenditure, steps, heart rate, sedentary duration, fair activity duration, and total active duration per day. In Wave 3 (mid-2019), Participant 169 registered better PRO for physical activity, depression and anxiety, memory, and sleep as well as more steps, a shorter sedentary duration, and longer light, fair, and vigorous durations. Social support was always high but never optimal. Nutrition and Quality of Life maintained high, but not optimal for waves 1 and 3. During the winter, the sleep duration was higher than during the summer. This real user example illustrates and emphasizes the importance of longitudinal state and behaviour assessments; we observed the

change of state in participant 169 as a change in the TechRO variables that indeed associated with worse PRO-based self-reported states.

Table 17. Summary of Characteristics of PRO (IPAQ, MSPSS, GADS, PREDIMED, SelfMNA, MFE, PSQI, EQ-5D-3L) and median TechRO (Fitbit) over the measurement period corresponding to each wave for Participant 169.

ID	Health	Wave	Country	Gender	Age	Physical Activity (IPAQ): Numeric Score	Social Support (MSPSS): Numeric Score	Depression and Anxiety (GADS): Numeric Score	Nutrition Mediterranean (PREDIMED): Numeric Score	Nutrition (SelfMNA): Numeric Score	Memory (MFE): Numeric Score	Sleep (PSQI): Numeric Score	Quality of Life (EQ-5D-3L): Health Score	TechRO: Energy (Kcal.)	TechRO: Steps (count)	TechRO: Heart Rate (bpm.)	TechRO: Sedentary (min.)	TechRO: Sedentary+Light (min.)	TechRO: Light (min.)	TechRO: Light+Fair (min.)	Fair (min.)	TechRO: Fair+Vigorous (min.)	TechRO: Vigorous (min.)	TechRO: Active (min.)	TechRO: Sleep (h:min.)
169	Diseased	1	Hungary	Female	69		5	68	7	12	13	15	80	2044.0	8035.0	52.5	842.0	999.0	192.5	253.0	23.0	51.0	19.0	300.0	7:06
169	Diseased	2	Hungary	Female	69	7338	5	51	5		15	15	75	1889.0	6076.0	56.0	843.0	994.0	207.0	245.0	21.0	51.5	22.5	273.5	7:08
169	Diseased	3	Hungary	Female	69	21,702	5	47	7	12	8	14	80	1979.0	8172.0	55.0	798.0	975.0	204.0	248.0	40.0	70.0	33.0	294.0	7:03
Median						14,520.0	5.0	51.0	7.0	12.0	13.0	15.0	80.0	1979.0	8035.0	55.0	842.0	994.0	204.0	248.0	23.0	51.5	22.5	294.0	7:06
Mean						14,520.0	5.0	55.3	6.3	12.0	12.0	14.7	78.3	1970.7	7427.7	54.5	827.7	989.3	201.2	248.7	28.0	57.5	24.8	289.2	7:05
SD						10,156.9	0	11.2	1.2	0.0	3.6	0.6	2.9	77.8	1172.6	1.8	25.7	12.7	7.7	4.0	10.4	10.8	7.3	13.9	0:03

Color coding: from orange (worse outcome) to yellow to green (better outcome).

4. Discussion

In this section we discuss our methodological approach (Section 4.1), the coQoL method in the perspective of past evidence (Section 4.2), observations on data quality (Section 4.3), and pathways towards personalized medicine (Section 4.4). We then review several limitations of our study (Section 4.5) and envision future work (Section 4.6).

4.1. Overall Methodological Approach in PROomics

The coQoL method explored patterns of correlations between PROs and TechROs towards their co-calibration. Consequently, we focused on identifying groups of strong correlations between PROs with a given recall period and TechROs, aggregating weeks to months of wearables data available before the administration day of the PRO. We considered correlations between similar latent constructs, e.g., PRO and TechRO physical activity or sleep, as high from 0.8 and above. However, for different latent constructs, such as PRO social support and TechRO sleep, where the probability of random correlation is low, correlations of even 0.5 are high. Hence, we presented in here correlations of 0.5 and above as of importance.

Due to the exploratory nature of our method, we deliberately omitted adjustments for multiple comparisons. The results of our method can guide future observational studies, as well as personalized, adaptive interventional studies, where the observational component will inform the intervention design *as we go*. Researchers can power such studies for enough confidence to exclude trivial effects.

4.2. coQoL in Perspective of Past Evidence

We recall that little prior research focused on assessing the relationships between sets of different outcomes assessed via PROs and consumer wearable TechROs in healthy seniors, in the wild, for extended periods (beyond the typical study duration of 7–14 days). On the one hand, past studies may have had similar to larger sample size, yet they have not yielded stronger statistical results; these co-calibrations rarely report values $r_S \geq 0.5$, as we do. On the other hand, we report a more

prolonged study duration (up to 2 years). The study duration of over a few weeks is essential to overcome the "novelty" effect of the technology (TechRO) on the state and behaviour of the user. Namely, the user, motivated by the feedback provided by the device while the study is being conducted, may move more or sleep differently, which then would be erroneously co-calibrated with the self-reports (PROs). The coQoL method leads to more accurate, real-world PRO- and TechRO-based datasets representing the real states and behaviours of the users. We define the past evidence in the context of *momentary co-calibration* efforts, where the PRO-TechRO co-calibrations may have been valid only for the short interval of data collection. Our proposed method coQoL expands the state of the art.

4.3. Observations on Data Quality

The wearable monitored some TechROs for more days than others. For example, the energy expenditure and steps appeared in most days. However, some days did not include durations of physical activity at increasing intensities, due to some seniors not wearing the wearable for enough hours that Fitbit recognized the activity or they did not reach the increased intensity physical activity on those days. Also, the TechROs that combine other TechROs, e.g., *fair+vigorous*, appeared in at most the minimum of the numbers of days when their constituent TechROs appeared. We acknowledge errors of a few days in long-term monitoring stemming from conditions beyond our control, such as errors at the device setup, at the recruitment site which took days to correct, or when running the automated data collectors from the seniors that were beyond our control in the project. These technological and human factors influenced the quality of the available data.

The wearable monitoring period may depend on the measured outcome, frequency of answers, and human factors. While the recall period of many scales is short (e.g., one week), collecting wearable data only for that duration may prove too strict. If the design is too strict, numerous participants will disqualify, and the results may bias in favour of diligent or adherent responders, who may also exhibit positive behaviours, e.g., exercising more diligently as well. Although some results indicate that 14–28 days of data could be enough for significant co-calibrations, the observations used in the co-calibration depend on the PRO answers and the TechRO data alike. If the participants are adherent to data collection for four weeks, but do not answer the questionnaire, the quality of the data may be insufficient to derive correlations. For some questionnaires, coQoL may relax the alignment (*leeway*) to account for human factors that contributed to data loss. On the other hand, a monitoring window of 120 days (4 months) may prove too wide to collect data reflecting the same behaviour as the reported one (the recall period), also because of the potential influence of seasonal effects. These seasonal, as well as other context dependencies, are illustrated when applying the coQoL to the MSPSS social support PRO. Our results indicate that having approximately one month of data before the administration of the MSPSS is sufficient to obtain significant correlations between *family trying to help* social support and *fair* activity even within a small sample of 39 participants. We observe that the MSPSS is time context-specific. Overall, across all questionnaires, we argue for an intermediary period of aggregation interval for TechRO not extending beyond 60–90 days.

4.4. Pathways towards Personalized Medicine

There is growing evidence within the medical domain that personal data paves a path towards personalized medicine, including genetics data and population-specific data, as well as, on a growing scale, data originating in the individuals' daily life environments and representing their natural, objective behaviours unfolding in different contexts of daily life. Daily life datasets are, in turn, collected via consumer wearables and smartphones with sensing capabilities.

From our study, we learn that an ideal wearable in the context of personalized medicine study would be comfortable to wear; should have a long battery life (at least a few days); should be accepted by individuals to use as their own, such that they forget they are in the study (implying minimal reactivity); and should provide relevant TechRO related to behavioural patterns (e.g., activity status, steps, as opposed to only heart rate, which would be hard to co-calibrate by itself).

Given our results, we also observe that for some PROs, different self-reported health status of the individuals yield different co-calibration results, even though our definition of disease refers only to mild self-reported cases. When the participants have a disease, other TechROs become correlated more strongly with other PROs than for the healthy ones. An observational study involving healthy individuals can leverage the coQoL method by monitoring a relevant subset of PRO/TechROs longitudinally, and occasionally co-calibrating the PROs with TechROs assuming the sensitivity of the coQoL method for when long-term, significant changes in TechRO occur. Based on the occasionally collected PRO answers, further in-depth examination of the individual's state may seek to understand if the TechRO change signals coincide with a significant and relevant PRO change, potentially implying a real change of the individual's health state. Once diagnosed, the individual's health state may be followed up, assuming another set of PRO/TechRO outcomes co-calibrated in time, to assess the change in the state of the disease accurately.

For example, in the case of diseased Participant 169, we observed that improvements or deteriorations in the state (as self-reported via the PROs for physical activity, Mediterranean diet, memory, and Quality of Life) coincided with TechROs (of physical activity in the sedentary, and light-vigorous spectrum, as well as the total physically active duration). Such trends are likely to differ between persons. As observed with Participant 169, administering the PROs only three times in two years and monitoring the TechRO behaviours using the wearable (minimally obtrusively, continuously, during daily life) yielded numerous trends across not only pairs of PROs and TechROs, but also across different PROs and TechROs.

The coQoL can provide a frontline approach to further triage the individual state assessment, for the healthy or diseased, without burdening the individuals with self-assessments, and at the same time without excluding participants who develop diseases and need to be monitored for long periods. In the context of the latter, the coQoL may be very suitable to assess changes of behaviour and health state in chronically ill patients.

We envision the following coQoL use case. The coQoL results can inform the design of longitudinal observations for selected individual PRO/TechRO outcomes, leveraged in personalized medicine solutions. The procedure consists of the observation for several consecutive days (for more TechRO-adherent participants, four weeks; for the less adherent participants, up to 3 months, from which one can derive around four weeks of quality data) followed by the co-calibration of TechROs with PROs. While monitoring, a potential gradual change in a subset of TechROs of interest can lead to contacting the individual for further health outcome assessments, via PRO or even clinical examination.

In new study designs, we suggest the study participation period of 60–90 days at most, and leverage behavioural techniques for participant wearable-adherence, to maximize the validity of the results acquired. The study design may imply repeated measures longitudinally over the years, e.g., PRO/TechRO co-calibration efforts over 60–90 consecutive days, repeated every few months up to a year (assuming same season every year).

4.5. Study Limitations

Several limitations characterize the presented here preliminary coQoL study. The first limitation is the small sample size, specific to an exploratory feasibility study. A second limitation is the resulting lack of power that reduced the complexity of the analysis method (i.e., statistical hypothesis tests). A third limitation is the presence of multiple PRO answers per individual for the same wave, albeit with high variability. However, we only included one answer per participant-wave to reduce bias towards diligent responders. In case of multiple answers per participant-wave, we chose the latest answer in time, to account for any form submission issues in the CoME software application or the participant changing their mind after submitting the answers once. A fourth limitation is a significant decrease in the number of participants data leveraged for the co-calibrations; we allowed for a *leeway* to allow PRO and TechRO alignments that are both (1) short-term, but accurate (e.g., 7–14 days, close to the

recall period), and (2) longitudinal, but permissive (e.g., 60–120 days, sufficient for the long-term behaviours to unfold). The study highlights the challenge of retaining individuals (shared by many health studies) that can provide outcomes through both self-report and a wearable that must be worn daily, over long periods.

4.6. Future Work

In the ongoing and future work, we expect to involve more participants for shorter periods (60–90 days), repeated every few months to a year, and focus on the PROs and TechROs delineated in this paper to deepen our knowledge about these specific co-calibration efforts and results. We plan to employ more advanced techniques and obtain more results within statistical significance as we increase the sample size in further studies aimed at calibrating PROs and TechROs for health outcomes and longitudinal behaviours such as physical activity and sleep in seniors. We aim to derive individual co-calibration trajectories models, as well as population models, e.g., similar groups of healthy or diseased individuals.

5. Conclusions

In this study, we present the coQoL method for co-calibrating the relationships between PROs and TechRO for eight PRO outcomes and TechRO behavioural markers of physical activity, sleep, and heart rate in a cohort of 42 seniors contributing data for two years. We reported human factors and quality properties from the data collected while their daily life unfolded. Our results can inform the design of personalized observational that assess daily life behaviours continuously and longitudinally, and that enable interventional studies towards reducing the risk of chronic disease and improve health and Quality of Life in the long term.

Author Contributions: Conceptualization, V.M. and K.W.; methodology, V.M. and K.W.; software, V.M.; validation, K.W.; formal analysis, V.M.; investigation, V.M. and K.W.; resources, K.W.; data curation, V.M. and K.W.; writing–original draft, V.M. and K.W.; writing–review and editing, V.M. and K.W.; visualization, V.M.; supervision, K.W.; project administration, K.W.; funding acquisition, K.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the H2020 Societal Challenges: 769765 (Wellbeing and Health Virtual Coach, *WellCo*), Active and Assisted Living programme: AAL-2014-7-127 (Caregivers and Me, *CoME*), and Active and Assisted Living programme: AAL-2019-6-120-CP (Social Robot Companion to Support Homecare Nurses, *Guardian*).

Acknowledgments: The authors thank Kedar K. V. Mate, Nancy E. Mayo, Kirke Kjellberg, Allan Berrocal Rojas, Robyn M. Stuart, Bo Markussen, Sofia Laghouila, Alexandre De Masi, and Matteo Ciman.

Conflicts of Interest: The authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application Programmable Interface
CLR	Centered Log Ratio
CLR PA	Centered Log Ratios of Physical Activity
CLR PA+S	Centered Log Ratios of Physical Activity and Sleep
CoME	Caregiver and Me
EQ-5D-3L	EuroQoL with 5 Domains and 3 Levels
GADS	Goldberg Anxiety and Depression Scale
IPAQ	International Physical Activity Questionnaire
MFE	Memory Failures of Everyday
MSPSS	Multidimensional Scale of Perceived Social Support
PREDIMED	Prevention with Mediterranean Diet
PRO	Patient-Reported Outcome

PSQI Pittsburgh Sleep Quality Index
QoL Quality of Life
SD Standard Deviation
SelfMNA Mini Nutritional Assessment
TechRO Technology-Reported Outcome

Appendix A. Literature Review

This section describes our procedure for literature review (Appendix A.1).

Appendix A.1. Literature Review Procedure

We searched for previous work by following a semi-structured approach, to prune papers distant from our research area from a vast body of literature. We agreed upon a hierarchy with properties divided into positive, neutral, and negative by their relative relevance to our research area (Figure A1).

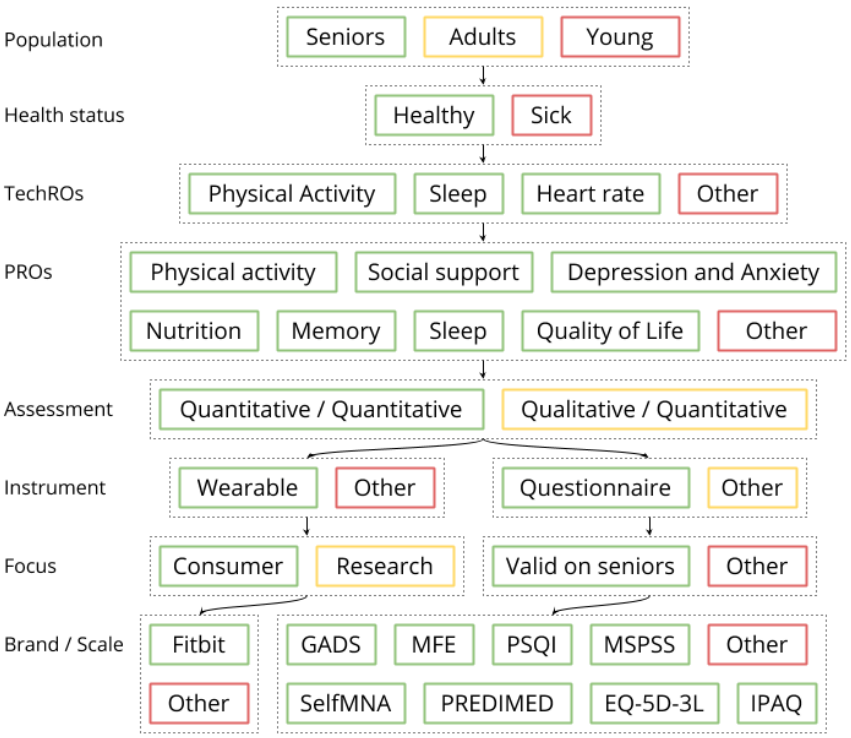


Figure A1. Related Work selection procedure (example on social support). Colors: green (positive towards inclusion), yellow (neutral), red (negative towards exclusion).

We began by including papers related to the PRO and using TechROs to the first level. We then followed a depth-first procedure of paper inclusion and exclusion. At each level, we included papers from the parent level and excluded all papers without positive properties for that level.

We then prioritized the papers by their deepest level of inclusion. We set the exclusion threshold at studies where the two outcomes, one PRO, and one TechRO, are used for co-calibration. We allowed only the PROs assessed in this paper (with a preference for the same questionnaires) and for TechROs provided by consumer wearables or accelerometers (with a preference for consumer wearables).

Numerous research directions and studies were excluded from our literature review reporting. We exclude papers that do not use PROs (or compare PROs) [72], do not use TechROs (or compare TechROs) [73], use other TechROs than wearables (e.g., smart phones [74], smart home [75], internet of things [76], medical imaging such as computer tomography or magnetic resonance [77]), focus on recognizing activities of daily life [78], or report only results following interventions [79].

Appendix B. Materials and Methods

In this section, we append notes on our materials and methods regarding patient-reported outcomes (Appendix B.1), technology-reported outcomes (Appendix B.2), and the co-calibration using coQoL (Appendix B.3).

Appendix B.1. Patient-Reported Outcomes (Questionnaires)

This part elaborates on our materials and methods for assessing the patient-reported outcomes: the used questionnaires (Appendix B.1.1), the administration of the questionnaires (Appendix B.1.2), the scoring of the answers (Appendix B.1.3), and the derivation of PRO variables (Appendix B.1.4).

Appendix B.1.1. Questionnaires

The participants provided PRO answers on questionnaires for physical activity (IPAQ [26]), social support (MSPSS [27]), anxiety and depression (GADS [28]), Mediterranean nutrition (PREDIMED [29,30]), nutrition (SelfMNA [31]), memory (MFE [32]), sleep (PSQI [33]), and health-related quality of life (EQ-5D-3L [34]). Table A1 illustrates the PRO questionnaires.

Table A1. Questionnaires with validated scales for PROs.

Outcome	Scale	Administration	Scoring
Profile	-	27 items assessing: age, gender, ethnicity, profession, education, cohabitants, height, weight, blood pressure, cholesterol, smoking, alcohol, medication (hypertension), personal health history (diabetes, apnea, insomnia, hyperglycemia, stroke, infarct, depression), and family health history (hypertension, diabetes, stroke, infarct, dementia)	-
Physical Activity	International Physical Activity Questionnaire (IPAQ) [26]	27 items of mixed types: yes/no, counts of days of physical activity per week, durations of physical activity per day. Recall: 2 weeks	Numeric score (estimated effort in metabolic equivalent of task). Categorical score with 3 levels: 0 low, 1 moderate, and 2 high. Numeric sub-scores for domains (work, leisure, transport, domestic and garden) and intensities of physical activity (sedentary, low, moderate, and vigorous).
Social Support	Multi-Dimensional Scale Perceived Social Support (MSPSS) [27]	12 items on a 7-level Likert scale (Q1–Q12). Recall: indefinite	Numeric score increasing with social support (1–2.9: low, 3–5: moderate, 5.1–7: high). Categorical score with 3 levels: 0 low, 1 moderate, and 2 high. Numeric sub-scores (1–7) for three sources of social support: significant other, family, and friends.
Anxiety and Depression	Goldberg depression and anxiety scale (GADS) [28]	18 items: 9 for Anxiety (denoted Q1A–Q9A), 9 for Depression (Q1D-Q9D), all on a 6-level Likert scale. The original answers were on a 2-level Likert scale. The collected answers are on a 6-level Likert scale. Recall: 1 month	Numeric score increasing with depression and anxiety: 0–9 no depression, 10–21 possible depression, 22–35 mild depression, 36–53 moderate depression, and 54–90 severe depression. Categorical score with 5 levels: 0 absent, 1 possible, 2 mild, 3 moderate, 4 severe.
Nutrition Mediterranean	Prevention with Mediterranean Diet (PREDIMED) [29,30]	14 binary items: 2 items yes/no, 12 items with thresholds for ingested food quantity (Q1–Q14). Recall: indefinite	Numeric score from 0–6 for no adherence to 7–12 for medium adherence to 13–14 for high adherence. Categorical score with 3 levels: 0 absent, 1 medium, 2 high.

Table A1. *Cont.*

Outcome	Scale	Administration	Scoring
Nutrition	Self-Reported Mini Nutritional Assessment (SelfMNA) [31]	6 items: 5 on various levels Likert scales, 1 binary (Q1–Q6). Recall: 3 months, same day	Numeric score from 0–7 for malnourished to 8–11 for risk of malnutrition to 12–14 for normal nutrition. Categorical score with 3 levels: 0 for malnutrition, 1 for risk, and 2 for normal nutrition.
Memory	Memory Failures of Everyday (MFE) [32]	28 items on a 3-level Likert scale (Q1–Q28). Recall: indefinite	Numeric score from 0 for no memory failures to 56 for potential memory failures. Categorical score separating 0 for no memory failures and 1 for potential memory failures, by comparing with deviations from the mean.
Sleep	Pittsburgh Sleep Quality Index (PSQI) [33]	25 items of mixed types: durations, yes/no, Likert scales (Q1, . . . , Q4, Q5A, . . . , Q5J, Q6, . . . , Q9). Recall: 1 month	Numeric score increasing as sleep quality decreases on a 0-21 scale. Categorical score of 1 for good sleep quality (0–4) and 0 for poor sleep quality (5–21). Numeric sub-scores (0–7) for: quality, latency, duration, efficiency, disturbance, medication, and daytime dysfunction.
Health-Related Quality of Life	EuroQoL health questionnaire (EQ-5D-3L) [34]	6 items: 5 on a 3-level Likert scale (denoted by their measured outcomes), 1 on a visual analog scale (Q1–Q6). Recall: same day	Numeric scores for five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, for the Likert items, increasing from 1 to 3 as life quality decreases. Visual analog scale of health state on the day of administration (giving a health score of 0–100), where higher numbers indicate better health.

Appendix B.1.2. Questionnaire Administration

For the participants in Spain, the partners used already available versions of the questionnaires in Spanish [80–87]. For the participants in Hungary, only some questionnaires had variants in Hungarian [88]. The local partners in the project translated the missing questionnaires from English to Hungarian (and assured the translation accuracy) to allow all participants to fill the PROs in their respective languages.

Appendix B.1.3. Answers Scoring

For the PRO questionnaires, we followed the scoring procedures set forth by the authors of the validated scales associated with each questionnaire. Only one questionnaire necessitated an additional assumption. For the physical activity questionnaire (IPAQ), we processed the individuals' physical activity answers by adhering to the data cleaning, maximum values for excluding outliers as described in the guide [89]. However, the guide does not provide a threshold for converting the *duration reported as weekly (not daily) to daily into an average daily time*. For example, if a senior reported seven hours of vigorous physical activity per day, the duration would likely reflect one hour per day. In this case, we allowed at most 7 h of physical activity per day at any intensity by dividing all excessive durations by 7 days.

Appendix B.1.4. Variables Derivation

We derived variables from both individual items, sub-scores, and scores of PRO scales. While the analysis of the scores exclusively would have been motivated by existing Rasch models providing calibrated positions of individual items and their sub-scores and scores [90], to our knowledge, there are no Rasch models for the PRO scales. Table 2 presents the derived PRO variables.

Appendix B.2. Technology-Reported Outcomes (Fitbit)

This part elaborates on our materials and methods for assessing the technology-reported outcomes: motivation and considerations for the Fitbit Charge 2 wearable (Appendix B.2.1), the processing of the wearable data (Appendix B.2.2), and the derivation of TechRO variables (Appendix B.2.3).

Appendix B.2.1. Fitbit Consumer Wearable

The space of consumer wearable manufacturers and devices is diverse, recording over 200 models [91], and the trend of adoption is increasing [13]. From all devices that provide physical activity and sleep TechROs, we chose Fitbit. Fitbit (1) monitors daily life behaviours accurately and continuously, (2) operationalizes the critical human factors for prolonged wear by senior end-users, and (3) facilitates reliable behavioural data collection.

First, Fitbit aims at motivating consumers to *"reach health and fitness goals by tracking activity, exercise, sleep, weight, and more"* [35]. It was selected for Digital Health software pre-certification by the US FDA [92]. Previous studies measured the accuracy of Fitbit consumer-friendly devices in reporting daily life behaviours of physical activity and sleep. For physical activity, Fitbit One and Zip had strong validity for step count and sleep duration, moderate for energy expenditure, and were weaker for fair and vigorous activity [12]. Fitbit Flex and Zip had adequate reliability and validity in measuring step count [93]. Fitbit Charge HR, Charge, Flex, Surge, Zip, and Alta agree with the ActiWatch GT3X+ research-grade accelerometer in assessing active minutes [37]. For sleep, Fitbit Charge HR can measure total sleep time [94] and time spent in bed [95] reliably, as compared with a sleep diary in a free-living setting or a research-grade accelerometer. For senior populations, Fitbit Charge 2 had better results in step count, energy expenditure, and sleep duration than the Garmin Vivosmart HR+ accelerometer in free-living environments [96]. Also, Fitbit One and Flex measure steps accurately in seniors [97].

Second, the positive senior user experience with the wearable is an essential factor that prolongs monitoring durations. For Fitbit, human factors studies found that over 90% of seniors agree that Fitbit was *"easy to use, useful, and acceptable"* over 8 months of wear [15] and seniors also place Fitbit the highest in usability (using the System Usability Scale [98]) among numerous other wearables [99]. Furthermore, the presence of a data display on the wristband leads to higher operation ratings [99].

Third, Fitbit provides a well-documented and developer-friendly application programming interface (API) which exposes a rich set of behavioural markers along [22] addressing goals of the project.

For our study, we selected the Fitbit Charge 2 wearable, a small wrist-worn watch which can monitor physical activity and sleep by using the same sensors such as those used in the validations, and displays steps, heart rate, and time, previously used in studies involving seniors (e.g., [96]).

Appendix B.2.2. Wearable Data Processing

To maintain high data quality, we considered *valid* days for the analysis only those days where the total duration of Fitbit monitoring was at least 21 h. We allowed at most three hours of missing data for device battery charging and handling (15–20 min to 2 h). Our choice reduced the impact of missing measurements and improved not only the measurement accuracy of TechRO behavioural markers in absolute daily durations but also enabled the assessment of TechRO behavioural markers relative to each other in the 24-h model of a day [64].

We constructed aggregate intervals with fixed durations of 7, 14, 21, 28, 60, 90, and 120 valid days to balance the number of included days in the analysis with the available intraday monitoring quality. The choice of 7 days for the lower bound was motivated by the need to acquire enough representative data for daily life, the 7 days as a common denominator of the PRO recall periods (where present), and the significant improvements in Fitbit accuracy for active minutes from 7 days onwards [37]. The choice of increasing intervals to the upper bound of 120 days reflected the duration of a wave, a large number of valid days per person (e.g., median 153 days for Spanish participants, Table A11), but also the high variance (a standard deviation of 113 days in Spain, Table A11).

We only included in the analysis intervals with at least 70% of their days valid, such that both weekdays and weekends were expected present in a week; the limit is compatible with previously reported consumer wearable use in seniors [100].

Appendix B.2.3. Variables Derivation

We split the TechROs into two amounts, *absolute* (behaviours in isolation, expressed in absolute amounts) and *relative* (behaviours relative to each other reflects the interdependences between behaviours during the 24 h of the day [64], expressed in relative amounts by the centred log ratios (CLR) of their compositions [65]).

In the absolute amount, we derived the variables into two families: *raw* and *processed*. We derived the raw daily energy expenditure (*energy*), step count (*steps*), and resting heart rate (*heart rate*) towards a total of 3 raw TechROs. We then derived the processed sedentary duration (*sedentary*), and the duration at three intensities (*light*, *moderate*, and *vigorous*) as processed by the Fitbit internal activity recognition algorithms. Since Fitbit had not published intensity thresholds, we also derived the cumulative durations in processed sedentary and light (*sedentary+light*), light and fair (*light+fair*), and fair and vigorous (*fair+vigorous*) intensities. We also calculated the total daily active duration (*active*) cumulating the light, fair, and vigorous processed durations. For sleep, we included the entire sleep duration of the day as a processed TechRO towards a total of 9 processed TechROs. We derived a total of 12 TechROs in the absolute amount.

For each aggregate interval duration and absolute TechRO, we used in the analysis as the aggregate the *median* from the absolute daily amounts as a variable. The 84 resulting variables are visible in the upper half of Table 3.

In the relative amount, we derived variables denoting compositional components of physical activity intensities and sleep throughout the day. We derived TechROs for each component of the centred log-ratio (*CLR*, [65]) transformation. The CLR is a symmetric transformation that does not require a reference component behaviour. We computed the CLRs of two families denoting distinct compositions: (1) from all physical activity durations (*CLR PA*) and (2) from all physical activity durations and the sleep duration (*CLR PA+S*), having 4 and 5 TechROs, respectively. We derived two relative families, as the CLRs of a composition do not translate to sub-compositions [65], but some studies may not be able to monitor sleep. We obtained a total of 9 TechROs in the relative amount.

For each aggregateinterval duration and relative TechRO, we used in the analysis as the aggregate the *geometric mean* from the relative daily amounts. The 63 resulting variables are visible in the lower half of Table 3.

The 147 derived TechRO variables can be seen in Table 3 (TechRO).

Appendix B.3. Co-Calibration Using coQoL

This part elaborates on our method coQoL to co-calibrate PROs and TechROs. The part covers the three types of analysis: descriptive (Appendix B.3.1), inferential (Appendix B.3.2), and pattern (Appendix B.3.3).

Appendix B.3.1. Descriptive Analysis (PROs and TechROs)

We describe the PROs and TechROs from two perspectives. The first perspective refers to the *values in the data*. The second perspective refers to the *amount of data*.

Within the first perspective, we describe the PROs by observing three summary statistics (median, mean, and standard deviation) of the participants-waves when grouped by health status (healthy vs. (mildly) diseased), country (Spain vs. Hungary), and gender (male vs. female) (Tables A3–A10).

Within the same perspective, we describe the TechROs by observing medians across the entire monitoring period (Table A12) in the first perspective.

Within the second perspective, we observe the counts of total and valid days (Table A11) within the same groups as for the first perspective.

Appendix B.3.2. Inferential Analysis (PROs vs. TechROs)

We set the leeway between PRO administration date and TechRO aggregate interval end date at (successively) 0, 7, 14, 21, 28, 60, 90, 120 days due to scarce exact matches. Pairs of variables with nearer such dates took precedence. We then analyzed lists of these pairs by using Spearman rank correlations. We chose this test as the best statistic to represent co-calibration motivated by the following assumptions. First, the PRO and TechRO variables were not independent (as they referred to the same participant). Second, the Spearman test is a nonparametric test that does not require an underlying distribution for the variables (some variables did not distribute normally, Shapiro Wilk normality test yielded $p < 0.05$-and some variables measured different metrics). Third, our aim was holistic in observing groups of significant correlations (and not individual correlations).

We only report the strongest correlation per TechRO interval duration. We consider correlations between distinct constructs (e.g., PRO social support and TechRO sleep duration) to be strong at $r_S \geq 0.5$ and associations between similar constructs (e.g., PRO and TechRO physical activity) to be strong at $r_S \geq 0.8$.

We consider a correlation coefficient significant when the extremities of its 95% confidence interval have the same sign. We avoided effect omissions at the expense of potential effects due to chance by not using adjustments for multiple tests [101] as our focus is on observing groups of correlations rather than individual correlations.

Appendix B.3.3. Pattern Analysis (PROs vs. TechROs)

For the pattern analysis, the *contour* metric separately counts for a significant and strong target correlation for a physical activity intensity (r_S 0.8 or above) the other significant correlations of the same sign at the lower and higher intensities. In case the intensity of the target correlation is at the extremity, the metric is undefined. In case the target correlation is adjacent to a correlation that has the opposite sign or is non-significant, the count on that side is 0. In case the correlation is unrelated to a physical activity intensity, this metric is undefined.

For example, the fair physical activity correlation 0.8 and the sequence of correlations [sedentary: 0.4*, sedentary+light: 0.5, light: 0.6*, light+fair: 0.6*, fair: 0.8*, fair+vigorous: 0.3*, and vigorous: −0.1*], where * denote significant correlations, has two correlations of lower intensities (0.6*, 0.6*) and one of higher intensity (0.3*). Figure A2 illustrates this case as Example (a). The figure contains three more examples.

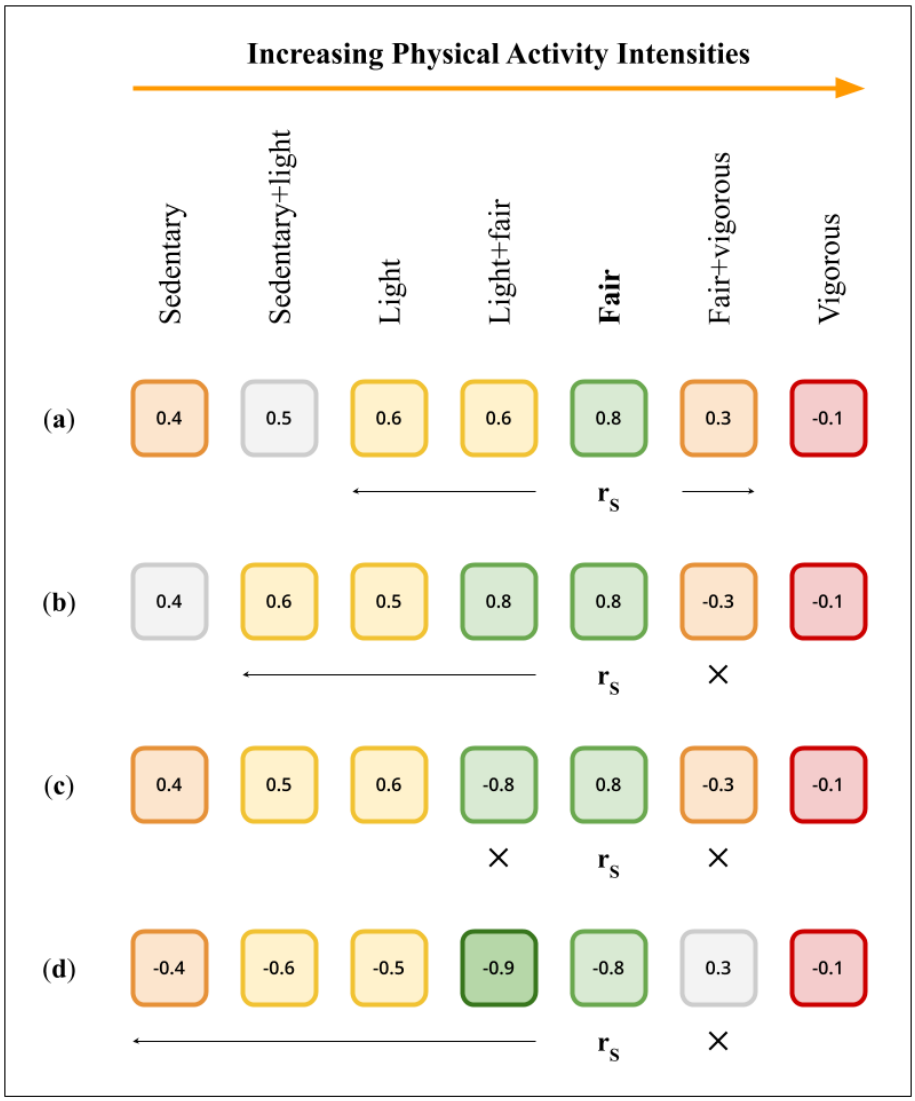


Figure A2. Examples of contours of correlations interrupted by non-significant or opposite-sign correlations. r_S marks the target correlation. × marks an interruption. Arrows mark the width of the contour. Only significant correlations are colored from red (weak) to green (strong). In example (a), the contour is interrupted by a non-significant correlation (at a lower intensity) and an opposite-sign correlation (at a higher intensity). Example (b) interrupts the entire right side of the contour by an opposite-sign correlation, represented with ×. Example (c) depicts a singleton contour, marked with × on both sides. Example (d) illustrates the rare case of a higher correlation than the target correlation, both in the same contour.

Appendix C. Results

This section includes results from our descriptive (Appendix C.1) and inferential analysis (Appendix C.2) analyses.

Appendix C.1. Descriptive Analysis (PROs and TechROs)

This part includes results from our descriptive analysis from patient-reported outcomes (Appendix C.1.1) and technology-reported outcomes (Appendix C.1.2).

Appendix C.1.1. Patient-Reported Outcomes (Questionnaires)

The 39 participants provided 289 answers (7.4 ± 4.4) on the 8 scales along the 3 waves. Table A2 depicts the numeric scores across waves.

Table A2. PRO numeric scores from answers by questionnaire ($N = 39$ participants).

PID	Health	Country	Gender	Age	Physical Activity (IPAQ)			Social Support (MSPSS)			Anxiety and Depression (GADS)			Mediterranean Nutrition (PREDIMED)			Nutrition (SelfMNA)			Memory (MFE)			Sleep (PSQI)			Quality of Life (IPAQ)		
					Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3	Wave 1	Wave 2	Wave 3
575	Healthy	Hungary	Female	65					5.0			8.0			5.0						7.0			3.0			85.0	
569	Healthy	Hungary	Female	67		23,238.0		7.0	6.0	6.0	19.0		11.0			7.0			11.0	8.0		6.0			4.0	80.0	80.0	95.0
133	Healthy	Hungary	Female	71		19,164.0	19,262.9	6.0	6.0	6.0	17.0		13.0			6.0	14.0			8.0		7.0	6.0		4.0	95.0	95.0	99.0
420	Healthy	Hungary	Female	71	576.0	2958.0		5.0	4.0			16.0			2.0			13.0			8.0			7.0		80.0	80.0	
215	Healthy	Hungary	Female	87	2446.0			5.0			33.0			3.0			10.0			8.0			12.0			80.0		
576	Healthy	Hungary	Male	60		2268.0			5.0																		95.0	
535	Healthy	Hungary	Male	69			8712.0			6.0			3.0			9.0			12.0			6.0			1.0			95.0
170	Healthy	Hungary	Male	70		8038.5	10,088.0		5.0	5.0			22.0			3.0						6.0			6.0		90.0	85.0
212	Healthy	Hungary	Male	72	8478.0	9793.5		5.0	4.0	5.0																90.0	90.0	90.0
419	Healthy	Hungary	Male	95		2016.0		5.0	4.0			13.0			6.0			12.0			12.0			7.0		90.0	95.0	
643	Healthy	Spain	Female	67		23,793.0		7.0	6.0		2.0	7.0			9.0		14.0			8.0	3.0		1.0	4.0		100.0	90.0	
798	Healthy	Spain	Female	67						6.0																		90.0
803	Healthy	Spain	Female	67																					5.0			80.0
617	Healthy	Spain	Female	69		3186.0		4.0	4.0			61.0			7.0			13.0			19.0			10.0		100.0	90.0	
620	Healthy	Spain	Female	69		3264.4			6.0			29.0			10.0			9.0			20.0			2.0			90.0	
640	Healthy	Spain	Female	69				5.0			26.0						14.0			10.0			6.0			70.0	
628	Healthy	Spain	Female	70				7.0			3.0						11.0			5.0			1.0			100.0	
638	Healthy	Spain	Female	71		6303.0		6.0	6.0		10.0						12.0			7.0			5.0			100.0	100.0	
648	Healthy	Spain	Female	72				6.0			5.0			11.0			13.0			3.0			1.0			80.0		
649	Healthy	Spain	Female	72																14.0						80.0		
795	Healthy	Spain	Female	72			2910.0			6.0			17.0			7.0			8.0			6.0			6.0			90.0
630	Healthy	Spain	Female	74				5.0			31.0						12.0			18.0			12.0			75.0		
411	Healthy	Spain	Male	45				5.0																		80.0		
790	Healthy	Spain	Male	66			10,101.0			6.0			3.0			11.0			14.0			7.0			4.0			100.0
700	Healthy	Spain	Male	67					3.0			10.0			8.0						7.0			9.0			40.0	
636	Healthy	Spain	Male	68		13,258.0		5.0	5.0																	40.0	40.0	
793	Healthy	Spain	Male	68			6560.0			5.0																		100.0
796	Healthy	Spain	Male	74			5907.0			4.0			2.0			9.0			14.0			0.0						80.0
502	Diseased	Hungary	Female	63																								80.0
169	Diseased	Hungary	Female	69		7338.0	21,702.0	5.0	5.0	5.0	68.0	51.0	47.0	7.0	5.0	7.0	12.0		12.0	13.0	15.0	8.0	15.0	15.0	14.0	80.0	75.0	80.0
132	Diseased	Hungary	Male	71				6.0	6.0	7.0	16.0		13.0			4.0				2.0	5.0	4.0			4.0	80.0	80.0	90.0
800	Diseased	Spain	Female	65						4.0			1.0			7.0			14.0			9.0			4.0			100.0
641	Diseased	Spain	Female	71		18,390.0		5.0	6.0		51.0	23.0			9.0		10.0			15.0	14.0			8.0		50.0	80.0	
624	Diseased	Spain	Female	72				6.0			21.0						14.0			5.0			8.0			80.0		
644	Diseased	Spain	Male	70				7.0																		40.0		
625	Diseased	Spain	Male	72				5.0			51.0						11.0			14.0			8.0			40.0		
634	Diseased	Spain	Male	72		15,748.5		5.0	5.0														3.0			40.0	40.0	
791	Diseased	Spain	Male	72			1953.0			7.0																		100.0
799	Diseased	Spain	Male	79						4.0			4.0			9.0			14.0			9.0			7.0			95.0

Color coding: from orange (worse score) to yellow to green (better).

Physical Activity (IPAQ)

We recorded 27 answers about physical activity on the IPAQ scale [26] that partitions physical activity into *low*, *moderate*, and *high* levels. The scale is described in depth in Appendix B.1.1. All participants recorded a median (mean ± SD) numeric score of 8038 (9535 ± 7106). There were 14 answers with a low categorical level of physical activity, one answer with a moderate level, and 12 answers with a high level. Table A3 enumerates the answers and Figure A3 depicts the sub-scores and scores by participant group.

Participant physical activity separated into two groups at the extremes of low and high physical activity. The levels only approximated the numeric scores, as the low categorical scores concentrated in the lower third of numeric scores and the high categorical scores concentrated in the upper third of numeric scores; the middle third included low and high levels of physical activity alike.

The participants from Hungary self-reported increased physical activity as compared to those from Spain, registering a median (mean ± SD) numeric score of 8478 (9738 ± 7370) compared to 6431 (9281 ± 6752) and a median categorical level of high physical activity compared to low physical activity.

Male participants reported increased levels of physical activity, registering a higher median numeric score of 8478 compared to 6820; however, the most active 5 participants contributed to a lower

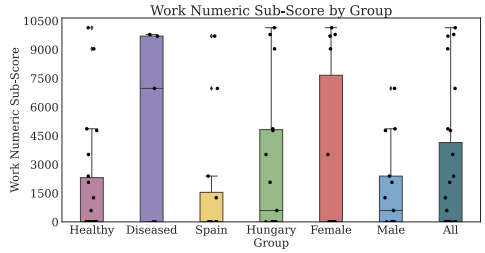
mean (SD) numeric score of 7916 (4038) compared to 11037 (8806) for the females. Woman participants registered higher variability in their self-reported physical activity than men.

Less than half (12/27) of the answers reported physical activity related to the work domain. Only a few (7/27) answers reported cycling as a means of transportation, and they associated with the upper half of numeric scores. The participants from Hungary reported increased physical activity as compared to those from Spain. Male participants reported increased median physical activity, and female participants reported increased mean physical activity.

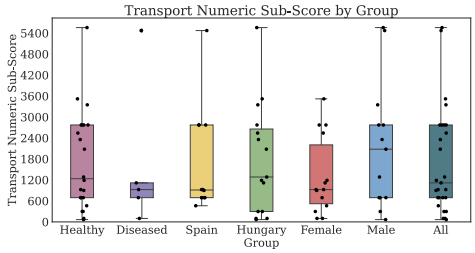
Table A3. Characteristics of PRO Physical Activity (IPAQ).

ID	Health	Wave	Country	Gender	Age	Work Domain Walking Minutes	Work Domain Moderate Minutes	Work Domain Vigorous Minutes	Work Domain Total	Active Transport Domain Walking Minutes	Active Transport Domain Cycling Minutes	Active Transport Domain Total	Domestic Home Domain Moderate Minutes	Domestic Garden Domain Moderate Minutes	Domestic Garden Domain Vigorous Minutes	Domestic Garden Domain Total	Leisure Domain Walking Minutes	Leisure Domain Moderate Minutes	Leisure Domain Vigorous Minutes	Leisure Domain Total	Work Domain Numeric Sub-Score	Leisure Domain Numeric Sub-Score	Active Transport Domain Numeric Sub-Score	Domestic Garden Domain Numeric Sub-Score	Numeric Score	Categorical Score
420	Healthy	1	Hungary	Female	71	0	0	0	0	90	0	90	60	0	0	60	30	0	0	30	0	99	297	180	576	0
791	Diseased	3	Spain	Male	72	0	0	0	0	210	0	210	420	0	0	420	0	0	0	0	0	0	693	1260	1953	0
419	Healthy	2	Hungary	Male	95	0	0	0	0	90	0	90	70	105	150	325	80	0	0	80	0	264	297	1455	2016	0
576	Healthy	2	Hungary	Male	60	20	360	70	450	20	0	20	10	10	0	20	20	0	0	20	2066	66	66	70	2268	1
215	Healthy	1	Hungary	Female	87	0	0	0	0	360	0	360	10	10	0	20	360	0	0	360	0	1188	1188	70	2446	0
795	Healthy	3	Spain	Female	72	0	0	0	0	140	0	140	420	0	0	420	360	0	0	360	0	1188	462	1260	2910	0
420	Healthy	2	Hungary	Female	71	0	0	0	0	30	0	30	920	0	0	920	30	0	0	30	0	99	99	2760	2958	0
617	Healthy	2	Spain	Female	69	0	0	0	0	280	0	280	210	210	0	420	240	0	0	240	0	792	924	1470	3186	0
620	Healthy	2	Spain	Female	69	0	0	0	0	273	0	273	360	61	0	421	315	0	0	315	0	1039	900	1324	3264	0
796	Healthy	3	Spain	Male	74	0	0	0	0	210	0	210	210	210	150	570	630	210	0	840	0	2919	693	2295	5907	0
638	Healthy	2	Spain	Female	71	0	0	0	0	210	0	210	420	840	0	1260	300	0	0	300	0	990	693	4620	6303	0
793	Healthy	3	Spain	Male	68	140	70	65	275	210	0	210	105	420	300	825	0	0	120	120	1262	960	693	3645	6560	2
169	Diseased	2	Hungary	Female	69	0	0	0	0	30	0	30	60	1680	0	1740	30	0	30	60	0	339	99	6900	7338	2
170	Healthy	2	Hungary	Male	70	540	350	210	1100	280	60	340	80	40	150	270	75	105	0	180	4862	667	1284	1225	8038	2
212	Healthy	1	Hungary	Male	72	0	0	0	0	360	360	720	180	180	0	360	300	240	240	780	0	3870	3348	1260	8478	0
535	Healthy	3	Hungary	Male	69	180	0	0	180	630	0	630	0	360	0	360	630	630	0	1260	594	4599	2079	1440	8712	2
212	Healthy	2	Hungary	Male	72	0	0	0	0	375	720	1095	0	60	0	60	120	0	450	570	0	3996	5557	240	9793	0
170	Healthy	3	Hungary	Male	70	210	420	300	930	350	200	550	140	140	300	580	100	0	0	100	4773	330	2355	2630	10,088	2
790	Healthy	3	Spain	Male	66	0	0	0	0	840	0	840	630	0	0	630	630	840	0	1470	0	5439	2772	1890	10,101	0
636	Healthy	2	Spain	Male	68	240	40	180	460	840	0	840	360	1440	60	1860	280	0	0	280	2392	924	2772	7170	13,258	2
634	Diseased	2	Spain	Male	72	840	840	105	1785	840	450	1290	120	180	30	330	315	225	15	555	6972	2059	5472	1245	15,748	2
641	Diseased	2	Spain	Female	71	2940	0	0	2940	280	0	280	840	840	0	1680	280	240	0	520	9702	1884	924	5880	18,390	2
133	Healthy	2	Hungary	Female	71	630	840	450	1920	630	240	870	420	420	240	1080	420	240	0	660	9039	2346	3519	4260	19,164	2
133	Healthy	3	Hungary	Female	71	540	1050	520	2110	840	0	840	420	360	171	951	420	150	90	660	10,142	2706	2772	3642	19,262	2
169	Diseased	3	Hungary	Female	69	420	420	840	1680	120	120	240	1260	1260	360	2880	0	0	0	0	9786	0	1116	10,800	21,702	2
569	Healthy	2	Hungary	Female	67	630	360	0	990	770	0	770	490	770	550	1810	910	550	550	2010	3519	9603	2541	7575	23,238	2
643	Healthy	2	Spain	Female	67	0	0	0	0	840	0	840	1470	1470	0	2940	1470	1470	0	2940	0	10,731	2772	10,290	23,793	0
Median: Healthy						0.0	0.0	0.0	0.0	315.0	0.0	350.0	210.0	160.0	0.0	495.5	300.0	0.0	0.0	337.5	0.0	1113.5	1236.0	1680.0	7299.0	0.0
Median: Diseased						420.0	0.0	0.0	1680.0	210.0	0.0	240.0	420.0	840.0	0.0	1680.0	30.0	0.0	0.0	60.0	6972.0	339.0	924.0	5880.0	15,748.0	2.0
Median: Spain						0.0	0.0	0.0	0.0	276.5	0.0	276.5	390.0	210.0	0.0	600.0	307.5	0.0	0.0	337.5	0.0	1113.5	912.0	2092.5	6431.5	0.0
Median: Hungary						20.0	0.0	0.0	180.0	350.0	0.0	360.0	80.0	140.0	0.0	360.0	100.0	0.0	0.0	180.0	594.0	667.0	1284.0	1455.0	8478.0	2.0
Median: Female						0.0	0.0	0.0	0.0	276.5	0.0	276.5	420.0	390.0	0.0	1015.5	307.5	0.0	0.0	337.5	0.0	1113.5	924.0	3951.0	6820.5	0.0
Median: Male						20.0	0.0	0.0	180.0	350.0	0.0	550.0	120.0	140.0	30.0	360.0	120.0	0.0	0.0	280.0	594.0	960.0	2079.0	1440.0	8478.0	1.0
Median: All						0.0	0.0	0.0	0.0	280.0	0.0	280.0	210.0	180.0	0.0	570.0	280.0	0.0	0.0	315.0	0.0	1039.0	1116.0	1890.0	8038.0	0.0
Mean: Healthy						142.2	158.6	81.5	382.5	394.0	71.8	465.8	317.5	323.0	94.1	734.6	350.9	201.5	65.9	618.4	1756.7	2491.5	1731.0	2762.3	8741.7	0.7
Mean: Diseased						840.0	252.0	189.0	1281.0	296.0	114.0	410.0	540.0	792.0	78.0	1410.0	125.0	93.0	9.0	227.0	5292.0	856.4	1660.8	5217.0	13,026.2	1.6
Mean: Spain						346.6	79.1	29.1	455.0	431.0	37.5	468.5	463.7	472.5	45.0	981.3	401.6	248.7	11.2	661.6	1694.0	2410.4	1647.5	3529.0	9281.0	0.6
Mean: Hungary						211.3	253.3	159.3	624.0	331.6	113.3	445.0	274.6	359.6	128.0	762.4	235.0	127.6	90.6	453.3	2985.4	2011.4	1774.4	2967.1	9738.4	1.1
Mean: Female						368.5	190.7	129.2	688.5	349.5	25.7	375.2	525.7	565.7	94.3	1185.8	368.9	189.2	47.8	606.0	3013.4	2357.4	1307.5	4359.3	11,037.8	0.8
Mean: Male						166.9	160.0	71.5	398.4	404.2	137.6	541.9	178.8	241.9	87.6	508.4	244.6	173.0	63.4	481.1	1763.1	2007.1	2160.0	1986.5	7916.9	1.0
Mean: All						271.4	175.9	101.4	548.8	375.8	79.6	455.4	358.7	409.8	91.1	859.7	309.0	181.4	55.3	545.9	2411.4	2188.7	1718.0	3216.8	9535.1	0.9
SD: Healthy						222.1	287.5	151.0	621.4	279.3	169.9	329.1	340.9	426.9	142.3	698.2	339.5	360.8	149.1	712.8	2918.6	2874.2	1384.6	2603.3	6816.5	0.9
SD: Diseased						1095.2	336.0	328.0	1135.5	284.7	174.3	448.2	453.7	634.2	141.4	947.5	141.7	114.0	12.0	254.7	4437.9	920.5	1936.0	3629.7	7299.5	0.8
SD: Spain						815.6	230.3	55.8	895.9	291.5	124.3	362.0	363.3	521.7	87.8	773.2	370.9	435.5	33.0	780.1	3103.9	2846.0	1445.5	2806.5	6752.6	0.9
SD: Hungary						253.0	324.5	249.8	748.7	267.8	195.0	348.6	362.0	487.6	164.9	799.1	259.3	200.6	172.8	549.7	3742.2	2564.1	1543.9	3086.5	7370.5	0.9
SD: Female						756.0	338.5	259.7	998.0	285.3	66.9	302.6	427.4	563.6	167.8	902.9	380.1	387.2	141.2	810.3	4307.0	3290.5	1077.2	3401.1	8806.6	0.9
SD: Male						246.3	249.2	96.1	534.8	277.5	223.1	385.3	179.7	368.3	108.1	444.1	234.8	259.6	130.7	466.3	2262.5	1849.5	1749.5	1745.9	4038.2	0.9
SD: All						579.4	299.3	200.6	821.7	282.9	171.5	354.8	374.6	506.1	142.2	795.1	324.5	332.0	136.5	670.1	3531.8	2700.3	1502.3	2978.4	7106.2	0.9

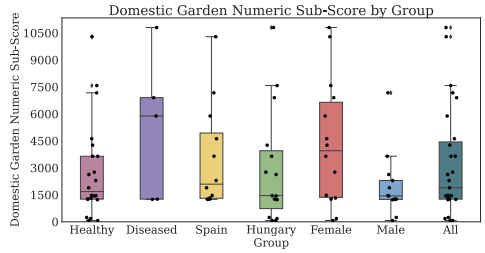
Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome).



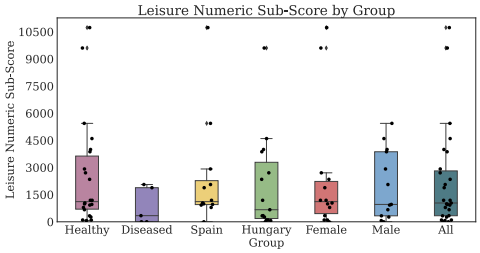
(a) Work Numeric Sub-Score



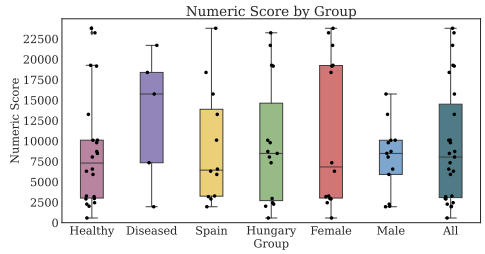
(b) Transport Numeric Sub-Score



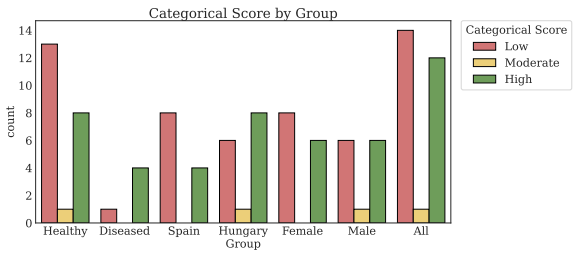
(c) Domestic and Garden Numeric Sub-Score



(d) Leisure Numeric Sub-Score



(e) Numeric Score



(f) Categorical Score

Figure A3. Sub-scores and Scores for Physical Activity (IPAQ).

Social Support (MSPSS)

Participants provided 55 answers on the MSPSS scale [27]. Their levels of social support were on a numeric scale from 1.0 to 7.0 corresponding to the categorical *low*, *moderate*, or *high* levels of social support. We describe this scale in Appendix B.1.1. All participants had a median (mean ± SD) numeric score of 5.0 (5.4 ± 0.9). Most answers corresponded to high social support. The levels of social support from separate sources (significant other, family, and friends) were also generally high. No answers reported low social support. Health status, country, and gender did not appear to change the level of social support fundamentally, neither by source nor in general. Table A4 enumerates the answers and Figure A4 depicts the sub-scores and scores by participant group.

Both healthy and diseased participants reported only slightly different levels of social support, as observed from the median (mean ± SD) of 5.0 (5.3 ± 0.9) healthy and 5.0 (5.5 ± 0.9) diseased. Participants with disease reported slightly higher significant other social support, registering mean numeric sub-scores of 5.8 compared to 5.5 for the significant other social support, 5.6 compared to 5.5 for the family social support, and 5.6 compared to 5.4 for the friends social support. Also, the answers had similar variations when comparing groups by health status. We observed no specific questions where the levels of social support differed by health.

Participants from Spain and Hungary self-reported similar levels of social support, registering similar medians (means) of 5.0 (5.4). Participants from Hungary self-reported more stable answers with SD 0.8 vs. 1.0.

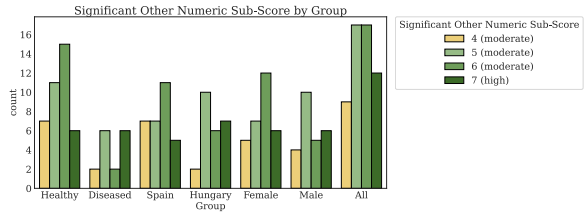
Table A4. Characteristics of PRO Social Support (MSPSS).

ID	Health	Wave	Country	Gender	Age	Q1: Special Person: Around When in Need	Q2: Special Person: Share Joys and Sorrows	Q3: Family: Tries to Help	Q4: Family: Gives Emotional Help and Support	Q5: Special Person: Real Source of Comfort	Q6: Friends: Try to Help	Q7: Friends: counted on when things go wrong	Q8: Family: Talk About Problems	Q9: Friends: Share My Joys and Sorrows	Q10: Special Person: Cares about Feelings	Q11: Family: Willing to Help Make Decisions	Q12: Friends: talk about problems	Significant Other Numeric Sub-Score	Family Numeric Sub-Score	Friends Numeric Sub-Score	Numeric Score	Categorical Score
700	Healthy	2	Spain	Male	67	5	5	4	2	2	3	4	2	6	4	4	5	4	3	4	3	1
420	Healthy	2	Hungary	Female	71	6	6	5	5	5	4	4	5	4	5	4	5	5	4	4	4	1
212	Healthy	2	Hungary	Male	72	3	5	5	6	5	5	5	5	5	5	5	5	4	5	5	4	1
419	Healthy	2	Hungary	Male	95	5	6	5	6	6	4	4	5	1	6	6	1	5	5	2	4	1
800	Diseased	3	Spain	Female	65	5	6	5	4	5	4	5	4	5	6	4	5	5	4	4	4	1
617	Healthy	1	Spain	Female	69	5	2	5	4	4	5	5	3	4	5	3	5	4	3	4	4	1
617	Healthy	2	Spain	Female	69	4	4	4	4	4	4	4	4	5	4	5	5	4	4	4	4	1
796	Healthy	3	Spain	Male	74	4	4	4	4	4	6	6	4	6	4	4	6	4	4	6	4	1
799	Diseased	3	Spain	Male	79	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	1
575	Healthy	2	Hungary	Female	65	5	6	7	7	3	5	5	6	6	5	3	5	4	5	5	5	1
169	Diseased	1	Hungary	Female	69	6	7	5	3	6	7	7	3	7	6	5	7	6	4	7	5	1
169	Diseased	2	Hungary	Female	69	5	5	6	5	5	6	7	5	7	6	6	7	5	5	6	5	1
169	Diseased	3	Hungary	Female	69	6	7	5	5	4	7	7	3	7	6	4	7	5	4	7	5	1
420	Healthy	1	Hungary	Female	71	6	6	6	5	6	1	5	5	5	6	4	5	6	5	4	5	1
215	Healthy	1	Hungary	Female	87	6	6	6	6	5	5	5	5	5	6	6	5	5	5	5	5	1
576	Healthy	2	Hungary	Male	60	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	1
170	Healthy	2	Hungary	Male	70	5	5	6	5	5	6	6	5	7	5	7	7	5	5	6	5	1
170	Healthy	3	Hungary	Male	70	5	5	6	4	6	6	6	5	6	5	5	6	5	5	6	5	1
212	Healthy	1	Hungary	Male	72	6	6	5	6	5	5	5	6	5	6	5	5	5	5	5	5	1
212	Healthy	3	Hungary	Male	72	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	1
419	Healthy	1	Hungary	Male	95	6	6	5	5	6	5	5	5	5	6	5	5	6	5	5	5	1
640	Healthy	1	Spain	Female	69	1	6	5	5	6	5	6	6	6	6	6	6	4	5	5	5	1
641	Diseased	1	Spain	Female	71	3	6	6	5	6	6	6	5	6	3	6	6	4	5	6	5	1
630	Healthy	1	Spain	Female	74	4	6	5	3	5	6	6	6	6	6	4	6	5	4	6	5	1
411	Healthy	1	Spain	Male	45	1	7	6	6	6	5	6	6	6	7	6	7	5	6	6	5	1
636	Healthy	1	Spain	Male	68	7	5	6	6	7	5	5	6	5	7	5	5	6	5	5	5	1
636	Healthy	2	Spain	Male	68	5	7	5	6	5	5	5	5	5	5	6	5	5	5	5	5	1
793	Healthy	3	Spain	Male	68	6	6	6	6	6	5	5	6	6	6	6	6	6	6	5	5	1
625	Diseased	1	Spain	Male	72	7	7	7	7	7	5	2	7	3	7	7	5	7	7	3	5	1
634	Diseased	1	Spain	Male	72	3	6	6	6	5	5	5	6	5	6	7	5	5	6	5	5	1
634	Diseased	2	Spain	Male	72	6	5	6	5	5	5	5	5	5	5	5	5	5	5	5	5	1
569	Healthy	2	Hungary	Female	67	6	7	7	7	7	5	5	7	5	7	7	5	6	7	5	6	2
569	Healthy	3	Hungary	Female	67	7	7	7	7	7	6	6	7	6	7	6	6	7	6	6	6	2
133	Healthy	1	Hungary	Female	71	7	7	7	7	7	6	6	7	6	7	6	7	7	6	6	6	2
133	Healthy	2	Hungary	Female	71	7	7	7	6	7	6	7	7	6	7	7	7	7	6	6	6	2
133	Healthy	3	Hungary	Female	71	6	7	7	7	7	6	6	7	6	7	6	6	6	6	6	6	2
535	Healthy	3	Hungary	Male	69	7	7	7	7	7	6	6	7	6	6	7	6	6	7	6	6	2
132	Diseased	1	Hungary	Male	71	7	7	7	7	7	6	6	6	5	7	6	5	7	6	5	6	2
132	Diseased	2	Hungary	Male	71	7	7	7	7	7	6	6	6	5	7	7	5	7	6	5	6	2
643	Healthy	2	Spain	Female	67	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	2
798	Healthy	3	Spain	Female	67	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	2
620	Healthy	2	Spain	Female	69	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	2
638	Healthy	1	Spain	Female	71	7	7	7	7	7	6	6	7	6	6	6	6	6	6	6	6	2
638	Healthy	2	Spain	Female	71	6	6	7	6	6	5	5	7	6	7	6	6	6	6	5	6	2
641	Diseased	2	Spain	Female	71	5	6	7	7	6	6	6	6	6	4	7	6	5	6	6	6	2
624	Diseased	1	Spain	Female	72	7	6	7	7	7	6	7	7	7	7	7	7	6	7	6	6	2
648	Healthy	1	Spain	Female	72	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	2
795	Healthy	3	Spain	Female	72	6	6	7	7	6	7	7	7	7	7	7	6	6	7	6	6	2
790	Healthy	3	Spain	Male	66	7	6	7	7	7	6	6	7	6	5	7	6	6	7	6	6	2
569	Healthy	1	Hungary	Female	67	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	2
132	Diseased	3	Hungary	Male	71	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	2
643	Healthy	1	Spain	Female	67	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	2
628	Healthy	1	Spain	Female	70	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	2
644	Diseased	1	Spain	Male	70	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	2
791	Diseased	3	Spain	Male	72	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	2
Median: Healthy						6.0	6.0	6.0	6.0	6.0	5.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.0	6.0	5.0	1.0
Median: Diseased						6.0	6.5	6.5	6.5	6.0	6.0	6.0	6.0	6.0	6.0	6.5	6.0	5.5	6.0	6.0	5.0	1.0
Median: Spain						6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.0	1.0
Median: Hungary						6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.0	6.0	6.0	6.0	5.0	6.0	5.0	5.0	5.0	1.0
Median: Female						6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	2.0
Median: Male						6.0	6.0	6.0	6.0	6.0	5.0	5.0	6.0	5.0	6.0	6.0	5.0	5.0	5.0	5.0	5.0	1.0
Median: All						6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.0	1.0
Mean: Healthy						5.5	5.9	5.9	5.7	5.7	5.3	5.5	5.7	5.6	5.9	5.5	5.6	5.5	5.4	5.3	5.3	1.4
Mean: Diseased						5.7	6.2	6.1	5.8	5.9	5.8	5.8	5.5	5.8	5.9	6.0	5.9	5.7	5.6	5.6	5.5	1.4
Mean: Spain						5.3	5.8	5.9	5.6	5.7	5.5	5.6	5.7	5.7	5.8	5.7	5.8	5.4	5.5	5.4	5.3	1.4
Mean: Hungary						5.9	6.2	6.0	5.8	5.8	5.4	5.7	5.6	5.5	6.0	5.6	5.6	5.7	5.4	5.4	5.3	1.4
Mean: Female						5.7	6.1	6.1	5.8	5.8	5.6	5.9	5.8	5.9	6.0	5.6	6.0	5.6	5.5	5.6	5.5	1.5
Mean: Male						5.4	5.8	5.8	5.7	5.7	5.3	5.3	5.5	5.3	5.8	5.7	5.4	5.5	5.5	5.2	5.1	1.2
Mean: All						5.6	6.0	6.0	5.7	5.8	5.5	5.6	5.6	5.6	5.9	5.7	5.7	5.5	5.5	5.4	5.3	1.4
SD: Healthy						1.4	1.0	0.9	1.2	1.1	1.1	0.8	1.1	1.0	0.9	1.1	1.0	0.9	1.0	0.9	0.9	0.4
SD: Diseased						1.3	0.9	0.9	1.3	1.0	0.9	1.3	1.3	1.2	1.2	1.1	1.0	1.0	1.1	1.2	0.9	0.4
SD: Spain						1.6	1.1	1.0	1.3	1.2	0.9	1.1	1.3	0.9	1.1	1.1	0.7	1.0	1.2	1.0	1.0	0.4
SD: Hungary						0.9	0.8	0.8	1.1	1.1	1.2	0.9	1.1	1.2	0.7	1.1	1.2	0.9	0.8	1.0	0.7	0.4
SD: Female						1.3	1.0	0.8	1.2	1.0	1.2	0.8	1.3	0.8	1.0	1.2	0.7	0.9	1.1	0.9	0.8	0.4
SD: Male						1.5	0.9	1.0	1.2	1.2	0.9	1.0	1.1	1.2	1.0	1.0	1.2	1.0	1.0	1.1	0.9	0.4
SD: All						1.4	1.0	0.9	1.2	1.1	1.1	1.0	1.2	1.1	1.0	1.1	1.0	1.0	1.0	1.0	0.9	0.4

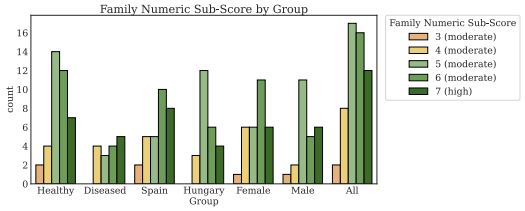
Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome).

Men self-reported lower social support than women, as observed in the median (mean ± std) numeric scores of 5.0 (5.2 ± 1.0) vs. 6.0 (5.5 ± 0.8) as well as median categorical score drop from high to moderate. Males self-reported less social support from the friends at means 5.2 vs. 5.6, less social

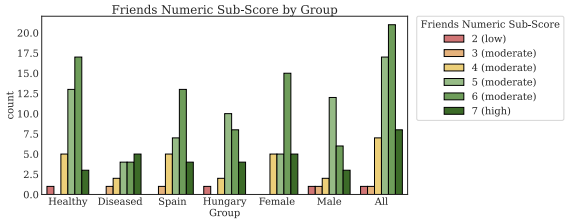
support from the significant other at means 5.5 vs. 5.6, and similar social support from the family at mean 5.5.



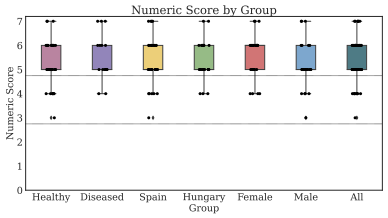
(a) Significant Other Numeric Sub-Score



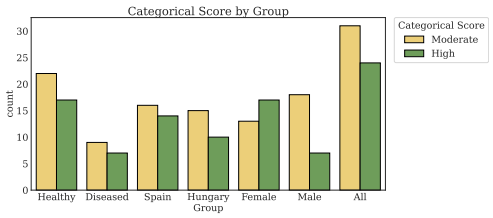
(b) Family Numeric Sub-Score



(c) Friends Numeric Sub-Score



(d) Numeric Score



(e) Categorical Score

Figure A4. Sub-scores and Scores for Social Support (MSPSS). Dotted markings delimit levels of the categorical score.

Anxiety and Depression (GADS)

We measured anxiety and depression through 34 answers on the GADS scale [28]. The scale assesses whether the anxiety and depression are categorized as *absent*, *possible*, *mild*, *moderate*, or *severe* through a numeric score from 0 to 90. It can be consulted in Appendix B.1.1. Participant mean ± SD numeric score was 20.8 ± 18.1. Participants self-reported absent anxiety and depression in 10 answers, possible anxiety and depression in 12 answers, mild in 6 answers, moderate in 4 answers, and severe in 2 answers. Table A5 enumerates the answers and Figure A5 illustrates the scores by participant group.

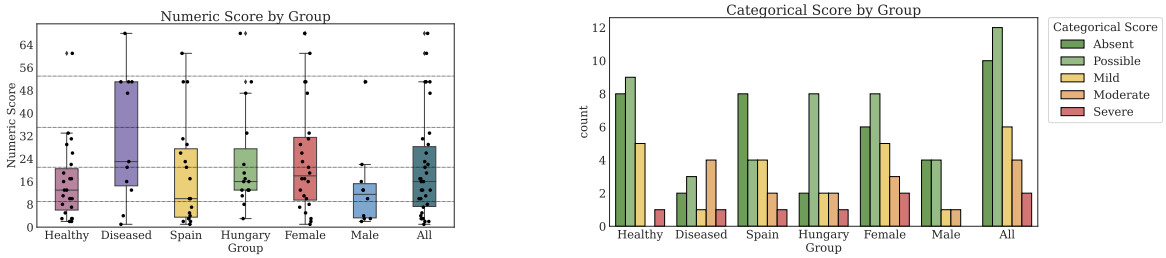
Most answers corresponding to moderate and severe anxiety and depression originated from participants who self-reported as diseased. Across the items and scores, the participants with disease reported more substantial anxiety and depression than the healthy participants, in particular for questions Q3A and Q7D. The median (mean ± SD) value for Q3A was 3.0 (2.0 ± 1.7) vs. 1.0 (0.9 ± 0.9). The median (mean ± SD) value for Q7D was 4.0 (2.8 ± 1.8) vs. 1.0 (1.3 ± 1.3), different by 2 and 3 levels, respectively. The median categorical scores were also different by one level, from possible to mild anxiety and depression. The answers from healthy participants had less variability than the answers from the participants with disease.

Across multiple items, women reported more anxiety and depression than male participants, yielding numeric scores higher by 8 units, as observed by the median (mean ± SD) scores of 18.0 (23.8 ± 18.8) compared to 11.5 (13.7 ± 13.9). They reported anxiety and depression with higher variability as well.

Table A5. Characteristics of PRO Anxiety and Depression (GADS).

ID	Health	Wave	Country	Gender	Age	Q1A: Keyed-up or on Edge	Q2A: Worrying a Lot	Q3A: Irritable	Q4A: Difficulty Relaxing	Q5A: Sleeping Poorly	Q6A: Headaches or Neck Aches	Q7A: Trembling [...] Urine	Q8A: Worried about Your Health	Q9A: Difficulty Falling Asleep	Q1D: Lacking Energy	Q2D: Lost Interest in Things	Q3D: Lost Confidence in Yourself	Q4D: Hopeless	Q5D: Difficulty Concentrating	Q6D: Lost Weight Due to Poor Appetite	Q7D: Waking Early	Q8D: Slowed Down	Q9D: Worse in the Mornings	Numeric Score	Categorical Score
800	Diseased	3	Spain	Female	65	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
643	Healthy	1	Spain	Female	67	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	2	0
796	Healthy	3	Spain	Male	74	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0
535	Healthy	3	Hungary	Male	69	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	3	0
628	Healthy	1	Spain	Female	70	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	3	0
790	Healthy	3	Spain	Male	66	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0
799	Diseased	3	Spain	Male	79	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0	0	0	0	4	0
648	Healthy	1	Spain	Female	72	0	0	0	0	1	0	1	1	0	1	0	0	0	0	0	0	1	0	5	0
643	Healthy	2	Spain	Female	67	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	7	0
575	Healthy	2	Hungary	Female	65	1	0	1	0	1	1	0	1	0	0	1	0	0	0	0	1	1	0	8	0
638	Healthy	1	Spain	Female	71	0	1	2	2	0	1	0	0	1	0	1	0	0	0	0	1	1	0	10	1
700	Healthy	2	Spain	Male	67	2	1	1	1	2	0	0	0	2	0	0	0	0	0	0	1	0	0	10	1
569	Healthy	3	Hungary	Female	67	1	1	1	0	2	1	1	0	2	0	0	1	0	0	0	0	0	1	11	1
133	Healthy	3	Hungary	Female	71	1	1	1	0	2	2	0	0	1	1	0	0	0	1	1	0	1	1	13	1
132	Diseased	3	Hungary	Male	71	1	1	1	0	1	1	0	1	0	1	1	1	0	0	0	1	2	1	13	1
419	Healthy	2	Hungary	Male	95	0	0	1	0	0	4	4	0	0	1	1	0	0	0	0	0	2	0	13	1
420	Healthy	2	Hungary	Female	71	1	1	0	0	1	3	0	0	1	0	1	2	0	1	1	0	4	0	16	1
132	Diseased	1	Hungary	Male	71	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	2	1	16	1
133	Healthy	1	Hungary	Female	71	1	0	1	2	1	1	1	0	1	1	2	0	0	1	1	2	2	0	17	1
795	Healthy	3	Spain	Female	72	1	1	0	1	2	0	0	1	2	2	1	0	0	0	4	1	1	0	17	1
569	Healthy	1	Hungary	Female	67	1	1	0	2	4	2	1	0	4	1	1	0	0	0	0	1	0	1	19	1
624	Diseased	1	Spain	Female	72	1	4	3	0	2	0	0	0	2	2	0	0	0	0	0	4	3	0	21	1
170	Healthy	3	Hungary	Male	70	1	2	1	1	2	1	1	1	0	1	1	2	1	1	1	3	2	0	22	2
641	Diseased	2	Spain	Female	71	2	4	3	2	1	2	0	0	0	2	0	0	0	1	0	4	1	1	23	2
640	Healthy	1	Spain	Female	69	1	1	1	1	1	4	2	3	2	3	0	0	0	4	0	1	2	0	26	2
620	Healthy	2	Spain	Female	69	2	2	2	1	3	4	0	2	2	1	2	0	0	1	0	3	4	0	29	2
630	Healthy	1	Spain	Female	74	4	3	2	2	2	2	0	3	1	3	0	0	0	1	2	4	2	0	31	2
215	Healthy	1	Hungary	Female	87	1	2	1	2	4	2	2	1	2	2	2	2	1	1	1	2	3	2	33	2
169	Diseased	3	Hungary	Female	69	4	1	4	3	4	0	1	1	4	4	4	5	1	1	1	4	4	1	47	3
169	Diseased	2	Hungary	Female	69	4	3	3	3	5	0	1	4	5	2	4	3	1	1	1	4	4	3	51	3
641	Diseased	1	Spain	Female	71	4	4	4	4	4	2	4	4	1	3	3	1	1	1	2	3	2	4	51	3
625	Diseased	1	Spain	Male	72	4	4	0	0	4	1	3	4	4	2	4	1	2	2	4	5	4	3	51	3
617	Healthy	2	Spain	Female	69	3	3	4	4	4	4	4	3	4	5	4	4	5	2	0	4	3	1	61	4
169	Diseased	1	Hungary	Female	69	4	1	4	4	5	4	3	3	5	4	4	5	4	3	2	5	5	3	68	4
Median: Healthy						1.0	1.0	1.0	0.0	1.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	13.0	1.0
Median: Diseased						2.0	1.0	3.0	1.0	2.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	0.0	1.0	1.0	4.0	2.0	1.0	23.0	2.0
Median: Spain						1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	10.0	1.0
Median: Hungary						1.0	1.0	1.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	2.0	1.0	16.0	1.0
Median: Female						1.0	1.0	1.0	1.0	2.0	1.0	0.0	0.5	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.5	1.5	0.0	18.0	1.0
Median: Male						1.0	0.5	0.5	0.0	0.5	1.0	0.5	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	11.5	1.0
Median: All						1.0	1.0	1.0	0.5	1.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.5	0.0	16.0	1.0
Mean: Healthy						1.0	1.0	0.9	0.8	1.3	1.4	0.7	0.6	1.0	1.0	0.8	0.4	0.3	0.5	0.5	1.3	1.2	0.2	15.6	1.0
Mean: Diseased						2.2	2.0	2.0	1.5	2.4	1.0	1.2	1.6	2.0	2.1	1.9	1.5	0.8	0.9	1.0	2.8	2.4	1.5	31.4	1.9
Mean: Spain						1.4	1.6	1.2	0.9	1.3	1.1	0.7	1.1	1.1	1.4	0.8	0.3	0.4	0.6	0.6	1.8	1.2	0.4	18.7	1.1
Mean: Hungary						1.4	0.9	1.3	1.2	2.2	1.6	1.0	0.8	1.7	1.2	1.5	1.4	0.5	0.7	0.6	1.6	2.1	0.9	23.3	1.4
Mean: Female						1.6	1.5	1.5	1.3	2.0	1.5	0.8	1.1	1.6	1.5	1.2	0.9	0.5	0.7	0.7	2.0	1.8	0.7	23.7	1.4
Mean: Male						1.0	0.9	0.5	0.3	1.0	0.9	1.0	0.7	0.7	0.9	1.0	0.5	0.3	0.4	0.6	1.3	1.2	0.5	13.7	0.9
Mean: All						1.4	1.3	1.2	1.0	1.7	1.3	0.9	1.0	1.3	1.3	1.1	0.8	0.4	0.6	0.6	1.7	1.6	0.6	20.7	1.2
SD: Healthy						0.9	0.9	0.9	1.0	1.3	1.4	1.1	1.0	1.2	1.2	1.0	1.0	1.0	0.9	0.9	1.2	1.2	0.5	13.4	0.9
SD: Diseased						1.6	1.6	1.7	1.6	1.8	1.2	1.3	1.6	2.0	1.1	1.7	1.8	1.1	0.8	1.2	1.8	1.6	1.3	21.6	1.3
SD: Spain						1.4	1.4	1.3	1.2	1.4	1.4	1.3	1.4	1.2	1.4	1.3	0.9	1.1	1.0	1.2	1.6	1.3	1.0	18.1	1.2
SD: Hungary						1.3	0.8	1.2	1.3	1.6	1.2	1.1	1.1	1.8	1.2	1.3	1.6	1.0	0.7	0.5	1.5	1.5	0.9	17.7	1.0
SD: Female						1.4	1.2	1.4	1.4	1.6	1.4	1.2	1.3	1.5	1.4	1.4	1.6	1.2	0.9	0.9	1.6	1.5	1.1	18.8	1.2
SD: Male						1.1	1.2	0.5	0.4	1.2	1.1	1.3	1.1	1.2	0.9	1.1	0.6	0.6	0.6	1.2	1.4	1.3	0.9	13.9	0.9
SD: All						1.3	1.3	1.3	1.3	1.6	1.3	1.2	1.3	1.5	1.3	1.4	1.4	1.1	0.9	1.0	1.6	1.5	1.0	18.1	1.1

Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome).



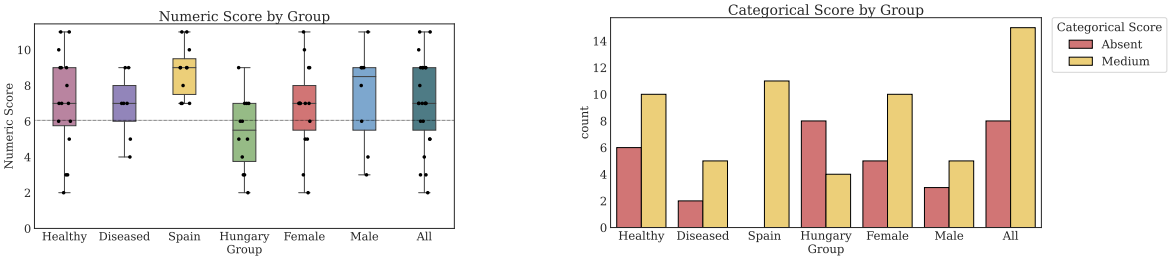
(a) Numeric Score **(b)** Categorical Score

Figure A5. Scores for Anxiety and Depression (GADS). Dotted markings delimit levels of the categorical score.

Mediterranean Nutrition (PREDIMED)

Participants self-reported their adherence to the Mediterranean diet by answering the PREDIMED scale [29,30] 23 times. The scale provides categorical scores for *absent*, *medium*, and *high* adherence using a numeric scale from 0 to 14 points, as described in Appendix B.1.1. Participants registered a mean ± SD numeric score of 7.0 ± 2.4. One-third of the answers corresponded to absent adherence to the Mediterranean diet, and two-thirds correspond to a medium adherence. Table A6 enumerates the answers. Figure A6 illustrates the scores by participant group.

A remarkable result is that among the nutrition diets none had high adherence to a Mediterranean diet. The scoring of the PREDIMED scale may explain this fact. It requires at least 13/14 items to be indicative of a Mediterranean diet to categorize the diet as highly adherent, while only 6/14 are necessary for medium adherence. The most adherent two participants only scored 11/14 and were thus categorized with medium adherence.



(a) Numeric Score **(b)** Categorical Score

Figure A6. Scores for Mediterranean Nutrition (PREDIMED). Dotted markings delimit levels of the categorical score.

One question that associated with the numeric and categorical scores is Q1 referring to olive oil as the primary culinary fat. Conversely, questions Q7 on sweet beverage use and Q13 on the preference for small animal meat had only 1/23 and 2/23 answers in the affirmative.

Participants from the healthy and diseased groups reported similar adherence, but higher variability, with means (SD) of 7.1 (2.7) and 6.9 (1.7), respectively.

The participant country of residence much coincided to the numeric score on the Mediterranean nutrition scale. All participants from Spain reported numeric scores of 7 or higher, corresponding to a medium adherence. Only one outlier person from Hungary had a numeric score of 9, and all other participants from Hungary had numeric scores of 7 or less. All participants categorized as having no adherence to the Mediterranean diet were from Hungary. Participants from Spain reported a median (mean ± SD) numeric score of 9.0 (8.8 ± 1.4) compared to 5.5 (5.3 ± 2.0) for Hungary. In general, the answers from the participants from Hungary had higher variance.

The answers from male participants indicated a higher adherence as depicted by the medians (means ± STD) of 8.5 (7.4 ± 2.6) and 7.0 (6.8 ± 2.3) on the numeric score, but also higher variability. However, there were fewer answers from men than women for this scale.

Table A6. Characteristics of PRO Mediterranean Nutrition (PREDIMED).

ID	Health	Wave	Country	Gender	Age	Q1: Olive Oil Main Culinary Fat	Q2: Olive Oil Use	Q3: Vegetables Use	Q4: Fruits Use	Q5: Red Meat, Hamburger, or Meat Use	Q6: Butter, Margarine, or Cream Use	Q7: Sweet/Carbonated Beverage Use	Q8: Wine Use	Q9: Legumes Use	Q10: Fish or Shellfish Use	Q11: Commercial Sweets or Pastries	Q12: Nuts Use	Q13: Preference to Small Animal Meat	Q14: Sofrito Use	Numeric Score	Categorical Score
420	Healthy	2	Hungary	Female	71	0	1.0	1.0	2.0	1.0	2.0	0.0	0.0	1.0	0.0	2.0	0.0	1.0	1.0	2	0
215	Healthy	1	Hungary	Female	87	0	0.0	1.0	2.0	1.0	2.0	0.0	0.0	1.0	0.0	2.0	3.0	0.0	2.0	3	0
170	Healthy	3	Hungary	Male	70	0	1.0	1.0	1.0	1.0	2.0	1.0	5.0	1.0	1.0	5.0	3.0	1.0	3.0	3	0
132	Diseased	3	Hungary	Male	71	0	0.0	2.1	1.1	1.0	2.0	0.0	1.0	2.1	1.0	2.0	1.2	1.0	2.0	4	0
575	Healthy	2	Hungary	Female	65	0	0.0	2.0	1.0	2.0	1.0	0.0	3.0	1.0	1.0	4.0	3.0	1.0	4.0	5	0
169	Diseased	2	Hungary	Female	69	0	0.0	1.0	2.0	0.3	1.0	0.0	2.0	0.5	0.5	1.0	0.3	1.0	4.0	5	0
133	Healthy	3	Hungary	Female	71	1	4.0	2.0	2.0	1.0	4.0	0.0	1.0	1.0	0.0	2.0	2.0	1.0	2.0	6	0
419	Healthy	2	Hungary	Male	95	1	1.0	2.0	2.0	1.0	2.0	0.0	1.0	1.0	0.0	0.0	0.0	1.0	2.0	6	0
569	Healthy	3	Hungary	Female	67	1	2.0	3.0	2.0	2.0	2.0	0.0	1.0	2.0	3.0	2.0	6.0	1.0	6.0	7	1
169	Diseased	1	Hungary	Female	69	1	3.0	2.0	1.0	0.2	1.0	0.0	5.0	2.0	0.2	0.0	0.0	1.0	5.0	7	1
169	Diseased	3	Hungary	Female	69	1	1.0	2.0	1.0	0.3	1.0	0.0	7.0	0.5	0.2	0.0	0.5	1.0	1.0	7	1
800	Diseased	3	Spain	Female	65	1	1.0	1.0	3.0	0.0	0.0	0.0	3.0	1.0	2.0	4.0	7.0	1.0	0.0	7	1
617	Healthy	2	Spain	Female	69	1	2.0	2.0	2.0	0.8	0.0	0.0	0.0	1.0	2.4	2.0	5.0	0.0	2.0	7	1
795	Healthy	3	Spain	Female	72	1	4.0	1.0	2.0	0.5	0.0	0.0	1.0	1.0	2.0	0.0	0.7	1.0	0.0	7	1
700	Healthy	2	Spain	Male	67	1	2.0	1.0	2.0	0.0	0.0	0.0	6.0	4.0	2.0	0.0	2.0	1.0	3.0	8	1
535	Healthy	3	Hungary	Male	69	1	2.0	2.0	5.3	0.1	0.0	0.0	3.0	2.0	1.0	3.0	3.0	1.0	3.0	9	1
643	Healthy	2	Spain	Female	67	1	2.0	2.0	4.0	0.5	0.0	0.0	7.0	2.0	2.0	0.0	1.0	1.0	1.0	9	1
641	Diseased	2	Spain	Female	71	1	3.0	2.0	2.0	0.0	0.0	0.0	2.0	1.0	7.0	7.0	4.0	1.0	7.0	9	1
796	Healthy	3	Spain	Male	74	1	3.0	2.0	3.0	0.0	0.0	0.0	1.0	2.0	2.0	0.0	2.0	1.0	3.0	9	1
799	Diseased	3	Spain	Male	79	1	1.0	1.0	3.0	0.0	0.0	0.0	3.0	2.0	2.0	0.0	4.0	1.0	2.0	9	1
620	Healthy	2	Spain	Female	69	1	2.0	2.0	4.0	1.0	0.0	0.0	1.0	2.0	3.0	0.0	4.0	1.0	2.0	10	1
648	Healthy	1	Spain	Female	72	1	2.0	2.0	3.0	0.8	0.0	0.0	2.0	6.0	3.0	7.0	5.0	1.0	5.0	11	1
790	Healthy	3	Spain	Male	66	1	10.0	3.0	4.0	1.0	0.0	0.0	0.0	5.0	1.0	0.0	7.0	1.0	4.0	11	1
Median: Healthy						1.0	2.0	2.0	2.0	1.0	0.0	0.0	1.0	1.5	1.5	2.0	3.0	1.0	2.5	7.0	1.0
Median: Diseased						1.0	1.0	2.0	2.0	0.2	1.0	0.0	3.0	1.0	1.0	1.0	1.2	1.0	2.0	7.0	1.0
Median: Spain						1.0	2.0	2.0	3.0	0.5	0.0	0.0	2.0	2.0	2.0	0.0	4.0	1.0	2.0	9.0	1.0
Median: Hungary						0.5	1.0	2.0	2.0	1.0	2.0	0.0	1.5	1.0	0.3	2.0	1.6	1.0	2.5	5.5	0.0
Median: Female						1.0	2.0	2.0	2.0	0.8	1.0	0.0	2.0	1.0	2.0	2.0	3.0	1.0	2.0	7.0	1.0
Median: Male						1.0	1.5	2.0	2.5	0.5	0.0	0.0	2.0	2.0	1.0	0.0	2.5	1.0	3.0	8.5	1.0
Median: All						1.0	2.0	2.0	2.0	0.8	0.0	0.0	2.0	1.0	1.0	2.0	3.0	1.0	2.0	7.0	1.0
Mean: Healthy						0.7	2.3	1.8	2.5	0.8	0.9	0.0	2.0	2.0	1.4	1.8	2.9	0.8	2.6	7.0	0.6
Mean: Diseased						0.7	1.2	1.5	1.8	0.2	0.7	0.0	3.2	1.3	1.8	2.0	2.4	1.0	3.0	6.8	0.7
Mean: Spain						1.0	2.9	1.7	2.9	0.4	0.0	0.0	2.3	2.4	2.5	1.8	3.7	0.9	2.6	8.8	1.0
Mean: Hungary						0.5	1.2	1.7	1.8	0.9	1.6	0.0	2.4	1.2	0.6	1.9	1.8	0.9	2.9	5.3	0.3
Mean: Female						0.7	1.8	1.7	2.2	0.7	0.9	0.0	2.3	1.5	1.7	2.2	2.7	0.8	2.8	6.8	0.6
Mean: Male						0.7	2.5	1.7	2.6	0.5	0.7	0.1	2.5	2.3	1.2	1.2	2.7	1.0	2.7	7.3	0.6
Mean: All						0.7	2.0	1.7	2.3	0.6	0.8	0.0	2.3	1.8	1.5	1.8	2.7	0.9	2.7	7.0	0.6
SD: Healthy						0.4	2.2	0.6	1.1	0.5	1.1	0.2	2.1	1.5	1.0	2.0	2.0	0.3	1.4	2.7	0.4
SD: Diseased						0.4	1.1	0.5	0.8	0.3	0.6	0.0	1.9	0.6	2.2	2.4	2.4	0.0	2.2	1.7	0.4
SD: Spain						0.0	2.3	0.6	0.7	0.4	0.0	0.0	2.1	1.6	1.4	2.7	2.0	0.2	2.0	1.4	0.0
SD: Hungary						0.5	1.2	0.5	1.1	0.5	0.9	0.2	2.1	0.5	0.8	1.4	1.7	0.2	1.4	1.9	0.4
SD: Female						0.4	1.2	0.5	0.9	0.5	1.1	0.0	2.2	1.2	1.8	2.2	2.2	0.3	2.1	2.3	0.4
SD: Male						0.4	2.9	0.6	1.3	0.4	0.9	0.3	2.0	1.3	0.6	1.7	1.9	0.0	0.6	2.5	0.4
SD: All						0.4	2.0	0.6	1.1	0.5	1.0	0.2	2.1	1.3	1.5	2.1	2.1	0.2	1.7	2.4	0.4

Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome).

Nutrition (SelfMNA)

We quantified participant nutrition through 24 self-reported answers on the SelfMNA scale [31]. The scale assesses a categorical nutrition status as *normal*, *at risk of malnutrition*, or *having malnutrition* and a numeric score between 0 and 14, as detailed in depth in Appendix B.1.1. Participants are well-nourished. Participants recorded a mean ± SD numeric score of 12.2 ± 1.7. More than two-thirds of the participants self-reported a healthy amount of nutrition, and the remaining answers reflected a

risk of malnutrition. One third obtained the maximum possible numeric score. None of the answers categorized the participant as malnourished. Table A7 depicts the answers and Figure A7 illustrates the scores by participant group.

Table A7. Characteristics of PRO Nutrition (SelfMNA).

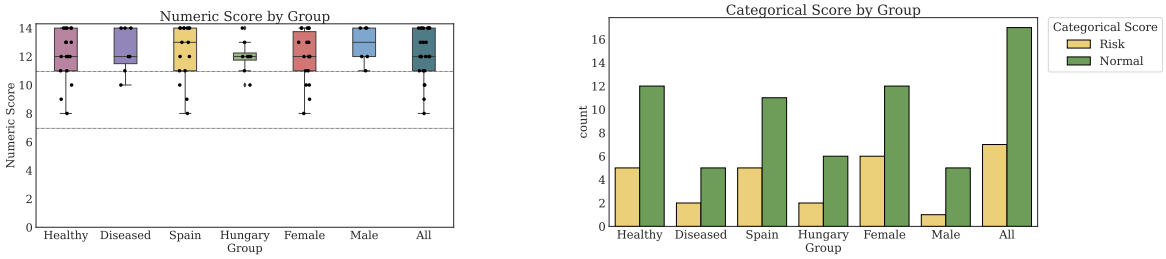
ID	Health	Wave	Country	Gender	Age	Q1: Food Intake Declined	Q2: Weight Lost	Q3: Described Current Mobility	Q4: Stressed or Severely Ill	Q5: Dementia and/or Severe Sadness	Numeric Score	Categorical Score
795	Healthy	3	Spain	Female	72	1	0	2	2	2	8	1
620	Healthy	2	Spain	Female	69	1	0	2	2	2	9	1
215	Healthy	1	Hungary	Female	87	2	3	1	0	1	10	1
641	Diseased	1	Spain	Female	71	2	3	2	0	2	10	1
569	Healthy	3	Hungary	Female	67	2	0	2	2	2	11	1
628	Healthy	1	Spain	Female	70	1	3	2	2	2	11	1
625	Diseased	1	Spain	Male	72	1	2	2	2	2	11	1
169	Diseased	1	Hungary	Female	69	2	3	2	0	2	12	2
169	Diseased	3	Hungary	Female	69	2	3	2	0	2	12	2
535	Healthy	3	Hungary	Male	69	1	2	2	2	2	12	2
419	Healthy	2	Hungary	Male	95	2	1	2	2	2	12	2
638	Healthy	1	Spain	Female	71	2	3	2	2	2	12	2
630	Healthy	1	Spain	Female	74	1	2	2	2	2	12	2
420	Healthy	2	Hungary	Female	71	1	3	2	2	2	13	2
617	Healthy	2	Spain	Female	69	1	3	2	2	2	13	2
648	Healthy	1	Spain	Female	72	1	3	2	2	2	13	2
133	Healthy	1	Hungary	Female	71	2	3	2	2	2	14	2
800	Diseased	3	Spain	Female	65	2	3	2	2	2	14	2
643	Healthy	1	Spain	Female	67	2	3	2	2	2	14	2
640	Healthy	1	Spain	Female	69	2	3	2	2	2	14	2
624	Diseased	1	Spain	Female	72	2	3	2	2	2	14	2
790	Healthy	3	Spain	Male	66	2	3	2	2	2	14	2
796	Healthy	3	Spain	Male	74	2	3	2	2	2	14	2
799	Diseased	3	Spain	Male	79	2	3	2	2	2	14	2
Median: Healthy						2.0	3.0	2.0	2.0	2.0	12.0	2.0
Median: Diseased						2.0	3.0	2.0	2.0	2.0	12.0	2.0
Median: Spain						2.0	3.0	2.0	2.0	2.0	13.0	2.0
Median: Hungary						2.0	3.0	2.0	2.0	2.0	12.0	2.0
Median: Female						2.0	3.0	2.0	2.0	2.0	12.0	2.0
Median: Male						2.0	2.5	2.0	2.0	2.0	13.0	2.0
Median: All						2.0	3.0	2.0	2.0	2.0	12.0	2.0
Mean: Healthy						1.5	2.2	1.9	1.8	1.9	12.1	1.7
Mean: Diseased						1.8	2.8	2.0	1.1	2.0	12.4	1.7
Mean: Spain						1.5	2.5	2.0	1.8	2.0	12.3	1.6
Mean: Hungary						1.7	2.2	1.8	1.2	1.8	12.0	1.7
Mean: Female						1.6	2.4	1.9	1.5	1.9	12.0	1.6
Mean: Male						1.6	2.3	2.0	2.0	2.0	12.8	1.8
Mean: All						1.6	2.4	1.9	1.6	1.9	12.2	1.7
SD: Healthy						0.4	1.1	0.2	0.4	0.2	1.7	0.4
SD: Diseased						0.3	0.3	0.0	0.9	0.0	1.4	0.4
SD: Spain						0.4	1.0	0.0	0.4	0.0	1.9	0.4
SD: Hungary						0.4	1.0	0.3	0.9	0.3	1.1	0.4
SD: Female						0.4	1.1	0.2	0.8	0.2	1.7	0.4
SD: Male						0.4	0.7	0.0	0.0	0.0	1.2	0.3
SD: All						0.4	1.0	0.1	0.7	0.1	1.7	0.4

Color coding: from orange (worse outcome relative to others) to yellow, to green (better outcome).

The groups of healthy and diseased participants were characterized by similar medians (12.0) and means (12.1 and 12.4), and only slight differences in the standard deviations (1.8 vs. 1.5). Healthy participants self-reported a decline in food intake for question Q1 while participants with disease reported being more stressed and severely ill in question Q4. Participants with disease reported less weight loss in Q2 as well as fewer variable answers across all items and scores except for Q4.

The participants from Spain reported similar levels of nutrition; however alternating ranks between questions: participants from Spain reported more decline in food intake in Q1, less weight loss in Q2, more mobility in Q3, and less stress, illness, dementia, or sadness in Q4 and Q5. Participants from Hungary reported had a more stable numeric score with a standard deviation of 1.11 for Hungary compared to 1.92 for Spain.

Women and men reported similar levels of nutrition, but provided more stable answers within their group, e.g., male standard deviation of 1.21 compared to female standard deviation of 1.79 for the numeric score.



(a) Numeric Score **(b)** Categorical Score

Figure A7. Scores for Nutrition (SelfMNA). Dotted markings delimit levels of the categorical score.

Memory (MFE)

Participants reported 36 answers on the MFE scale for memory [32]. The scale classifies memory failures as *absent* or *potential* through a numeric score from 0 to 56. See the description of MFE in Appendix B.1.1. Participants had mean ± SD numeric score of 8.7 ± 4.7. The median and mean numeric scores indicate absent memory failures. One-third of the answers indicate the possibility of memory failures, originating predominantly from female participants from Spain. Table A8 enumerates the answers. Figure A8 illustrates the scores by participant group.

One item whose answers may associate with the numeric score is Q15: Forgetting important details of done things.

The participants self-reported as diseased reported a higher probability of memory failures, as seen in the median (mean ± SD) numeric score of 9 (9.41 ± 4.5) compared to 7 (8.45 ± 4.8) for healthy participants. The ranking for the medians and means for individual items between the healthy and diseased alternate. Examples of questions where the diseased fared worse include Q5 (checking whether something was done), Q6 (forgetting time of events), Q14 (forgetting to do planned things), and Q18 (forgetting to tell somebody something important) as seen from the medians different by 1 out of the maximum two levels as well as the slightly different means. Healthy and diseased participants had similar variability in the numeric scores and alternating ranks of variability within individual questions.

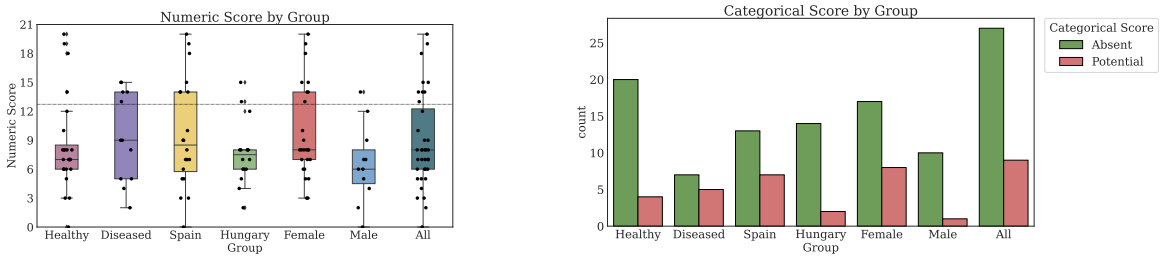
The participants from Hungary may have slightly fewer chances of memory failure, as observed from the medians (means) of 7.5 (7.7) and 8.5 (9.7) different by 1 (2) points. Furthermore, the numeric scores from the participants from Hungary are more stable. Questions Q5 (checking whether something was done) and Q6 (forgetting time of events) indicate the potential memory decline within the subjects from Spain. Question Q8 (being reminded about things) indicates the opposite. Other questions that weigh towards an expected increase in memory failures for the participants from Spain are Q7 (being reminded about things), Q21 (telling someone a story or joke repeatedly), and Q24 (forgetting where things are normally kept).

Men self-reported improved memory numeric scores as compared to women, as seen from the medians (means) of 6 (6.54) and 8 (9.76), respectively. Questions that contribute to this difference are Q6, Q8, and Q24 and against this difference Q5. Males self-reported more stable memory failures, as seen from the SD 3.86 and SD 4.76, respectively.

Table A8. Characteristics of PRO Memory (MFE).

ID	Health	Wave	Country	Gender	Age	Q1: Forgetting Objects put	Q2: Failing to Recognise Places	Q3: Finding a Television Story Difficult	Q4: Not Remembering a Change in Daily Routine	Q5: Checking Whether Something Was Done	Q6: Forgetting Time of Events	Q7: Completely Forgetting to Take Things	Q8: Being Reminded about Things	Q9: Reading Anew Already Read Something	Q10: Letting Ramble about Unimportant Things	Q11: Failing to Recognise Close Relatives or Friends	Q12: Having Difficulty Picking up a New Skill	Q13: Finding Word Is "on the tip of the Tongue"	Q14: Forgetting Forgetting to do Planned Things	Q15: Forgetting important details of done things	Q16: Forgetting the Topic of an Ongoing Conversation	Q17: Failing to follow a Story in a Newspaper	Q18: Forgetting to tEll Somebody Something Important	Q19: Forgetting Important Details about Oneself	Q20: Getting Told Details Mixed up and Confused	Q21: Telling Someone a Story or Joke Repeatedly	Q22: Forgetting Details of Things You Do Regularly	Q23: Finding Famous Faces Unfamiliar	Q24: Forgetting Where Things Are Normally Kept	Q25: Getting Lost Where You Have OFTEN been before	Q26: Getting Lost Where You Have Been RARELY before	Q27: Doing Some Routine Thing Twice by Mistake	Q28: Repeating to sOmeone What You Have just Told Them	Numeric Score	Categorical Score
796	Healthy	3	Spain	Male	74	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
132	Diseased	1	Hungary	Male	71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2	0
643	Healthy	2	Spain	Female	67	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0
648	Healthy	1	Spain	Female	72	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3	0
132	Diseased	3	Hungary	Male	71	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	4	0
132	Diseased	2	Hungary	Male	71	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	1	0	0	5	0
628	Healthy	1	Spain	Female	70	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	5	0
624	Diseased	1	Spain	Female	72	1	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5	0
569	Healthy	3	Hungary	Female	67	0	0	0	0	0	0	0	1	0	1	0	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	6	0
535	Healthy	3	Hungary	Male	69	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	6	0
170	Healthy	3	Hungary	Male	70	1	1	0	1	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0
795	Healthy	3	Spain	Female	72	1	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	6	0
575	Healthy	2	Hungary	Female	65	2	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	7	0
133	Healthy	3	Hungary	Female	71	0	0	1	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	7	0
638	Healthy	1	Spain	Female	71	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	7	0
790	Healthy	3	Spain	Male	66	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	1	1	0	7	0
700	Healthy	2	Spain	Male	67	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	7	0
569	Healthy	1	Hungary	Female	67	0	0	0	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	8	0
169	Diseased	3	Hungary	Female	69	1	0	0	0	0	1	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	8	0
133	Healthy	1	Hungary	Female	71	0	0	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	8	0
420	Healthy	2	Hungary	Female	71	1	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	8	0
215	Healthy	1	Hungary	Female	87	0	0	0	0	0	1	0	1	0	0	0	0	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	8	0
643	Healthy	1	Spain	Female	67	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	8	0
800	Diseased	3	Spain	Female	65	1	0	0	0	1	1	1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	9	0
799	Diseased	3	Spain	Male	79	0	0	0	0	1	1	1	1	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	9	0
640	Healthy	1	Spain	Female	69	1	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	1	1	1	0	1	0	0	0	0	0	10	0
419	Healthy	2	Hungary	Male	95	2	0	0	1	1	1	1	1	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	12	0
169	Diseased	1	Hungary	Female	69	2	0	0	0	0	0	1	1	0	0	1	0	2	1	1	1	0	1	0	0	1	1	0	0	0	0	0	0	0	13	1
641	Diseased	2	Spain	Female	71	0	0	2	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	14	1
649	Healthy	1	Spain	Female	72	1	0	1	0	0	1	0	0	1	1	0	1	1	1	0	0	0	1	0	0	1	0	2	1	0	1	0	0	14	1
625	Diseased	1	Spain	Male	72	1	0	0	0	1	1	1	1	0	0	0	0	1	1	1	0	0	1	0	0	1	1	1	1	0	0	0	1	14	1
169	Diseased	2	Hungary	Female	69	2	0	1	0	1	1	0	1	1	1	0	0	1	1	1	1	0	1	0	0	1	0	0	0	0	1	0	0	15	1
641	Diseased	1	Spain	Female	71	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1	0	15	1
630	Healthy	1	Spain	Female	74	1	1	0	0	1	1	2	1	1	0	0	1	1	0	0	0	1	0	0	0	1	1	2	1	0	0	1	1	18	1
617	Healthy	2	Spain	Female	69	1	0	0	0	1	1	1	1	0	1	0	1	2	1	1	0	0	1	0	1	1	0	1	2	0	0	1	1	19	1
620	Healthy	2	Spain	Female	69	1	0	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	0	1	1	1	0	1	1	1	20	1
Median: Healthy						1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.5	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	7.0	0.0
Median: Diseased						0.5	0.0	0.0	0.0	1.0	1.0	0.5	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.5	0.0	0.0	9.0	0.0
Median: Spain						1.0	0.0	0.0	0.0	1.0	1.0	1.0	0.5	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	1.0	0.0	0.0	0.0	0.0	8.5	0.0
Median: Hungary						0.5	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.5	0.0
Median: Female						1.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	8.0	0.0
Median: Male						0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	0.0
Median: All						1.0	0.0	0.0	0.0	0.5	0.5	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	8.0	0.0
Mean: Healthy						0.7	0.0	0.1	0.1	0.4	0.4	0.4	0.5	0.1	0.2	0.0	0.5	0.8	0.2	0.1	0.1	0.0	0.3	0.0	0.2	0.4	0.2	0.4	0.5	0.0	0.2	0.2	0.2	8.4	0.1
Mean: Diseased						0.6	0.0	0.2	0.1	0.5	0.6	0.5	0.6	0.3	0.0	0.0	0.1	0.9	0.5	0.4	0.3	0.0	0.5	0.0	0.0	0.3	0.1	0.1	0.5	0.2	0.5	0.2	0.0	9.4	0.4
Mean: Spain						0.7	0.0	0.2	0.1	0.6	0.6	0.6	0.5	0.3	0.1	0.0	0.4	1.0	0.3	0.2	0.0	0.1	0.4	0.0	0.1	0.5	0.2	0.5	0.6	0.0	0.3	0.4	0.3	9.6	0.3
Mean: Hungary						0.7	0.0	0.1	0.1	0.3	0.3	0.3	0.7	0.1	0.2	0.0	0.3	0.6	0.4	0.2	0.3	0.0	0.3	0.0	0.2	0.3	0.1	0.1	0.4	0.1	0.3	0.0	0.1	7.6	0.1
Mean: Female						0.8	0.0	0.2	0.1	0.4	0.5	0.5	0.6	0.3	0.2	0.0	0.4	1.0	0.4	0.2	0.2	0.1	0.4	0.0	0.2	0.4	0.2	0.3	0.6	0.0	0.2	0.2	0.2	9.7	0.3
Mean: Male						0.5	0.0	0.0	0.1	0.5	0.4	0.3	0.4	0.0	0.0	0.0	0.1	0.5	0.3	0.1	0.0	0.0	0.2	0.0	0.0	0.3	0.0	0.3	0.2	0.2	0.3	0.1	0.1	6.5	0.0
Mean: All						0.7	0.0	0.1	0.1	0.5	0.5	0.4	0.6	0.2	0.1	0.0	0.3	0.8	0.3	0.2	0.1	0.0	0.4	0.0	0.1	0.4	0.1	0.3	0.5	0.0	0.3	0.2	0.2	8.7	0.2
SD: Healthy						0.5	0.2	0.3	0.3	0.4	0.4	0.5	0.4	0.3	0.4	0.0	0.5	0.4	0.4	0.3	0.3	0.2	0.4	0.0	0.4	0.4	0.4	0.6	0.5	0.0	0.4	0.4	0.4	4.8	0.3
SD: Diseased						0.7	0.0	0.5	0.3	0.4	0.4	0.5	0.4	0.4	0.2	0.2	0.3	0.6	0.4	0.4	0.4	0.2	0.4	0.0	0.2	0.4	0.3	0.3	0.5	0.4	0.5	0.4	0.2	4.4	0.4
SD: Spain						0.4	0.2	0.5	0.3	0.4	0.4	0.5	0.5	0.4	0.3	0.0	0.4	0.3	0.4	0.4	0.2	0.3	0.4	0.0	0.3	0.5	0.4	0.6	0.5	0.0	0.4	0.4	0.4	5.5	0.4
SD: Hungary						0.8	0.2	0.3	0.3	0.4	0.4	0.4	0.4	0.3	0.4	0.2	0.4	0.5	0.4	0.4	0.4	0.0	0.4	0.0	0.4	0.4	0.3	0.3	0.4	0.3	0.4	0.0	0.3	3.1	0.3
SD: Female						0.6	0.1	0.5	0.3	0.4	0.4	0.5	0.4	0.4	0.4	0.1	0.4	0.4	0.4	0.4	0.4	0.3	0.4	0.0	0.4	0.4	0.4	0.6	0.5	0.0	0.4	0.4	0.4	4.7	0.4
SD: Male						0.6	0.2	0.0	0.3	0.4	0.4	0.4	0.4	0.0	0.2	0.0	0.3	0.4	0.4	0.3	0.2	0.0	0.4	0.0	0.2	0.4	0.2	0.4	0.4	0.4	0.4	0.3	0.3	3.8	0.2
SD: All						0.6	0.2	0.4	0.3	0.5	0.5	0.5	0.4	0.4	0.3	0.1	0.4	0.5	0.4	0.4	0.3	0.2	0.4	0.0	0.3	0.4	0.3	0.5	0.5	0.2	0.4	0.4	0.4	4.7	0.4

Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome).



(a) Numeric Score **(b)** Categorical Score

Figure A8. Scores for Memory (MFE). Dotted markings delimit levels of the categorical score.

Sleep (PSQI)

The seniors self-reported their sleep quality through 32 answers on the PSQI scale [33]. PSQI assesses sleep quality as *good* or *poor* based on a numeric score from 0 to 21, as described in Appendix B.1.1. Participants recorded a median (mean ± SD) numeric score of 6.0 (6.3 ± 3.9). The median and mean sleep quality situated at the better extremity of poor sleep quality. Two-fifths of the answers corresponded to poor sleep quality. Table A9 enumerates the answers. Figure A9 illustrates the sub-scores and scores by participant group.

The participants with disease self-reported less adequate sleep, as depicted by the median (mean ± SD) of 8.0 (8.6 ± 3.2) compared to 5.0 (5.3 ± 4.3). Participants with disease self-reported less adequate sleep through questions Q5B (trouble sleeping due to waking up in the middle of the night) with a difference between median (mean) answers of 1.5 (0.53) out of 3. Conversely, healthy participants self-reported decreased sleep quality due to using the bathroom in Q5C with a median (mean) difference of 1.0 (0.55) out of 3. The healthy participants provided more stable PROs with a standard deviation for the numeric score of 3.23 as compared to 4.34.

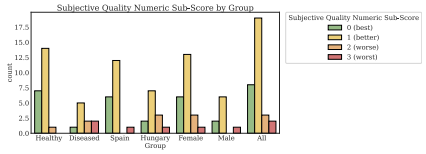
The participants from Hungary reported worse sleep quality with a median (mean ± SD) of 6.0 (7.5 ± 0.2) in Hungary compared to 5.0 (5.5 ± 0.1) in Spain. The difference between the sleep quality for participants in Hungary and Spain is visible in the numeric sub-scores, e.g., subjective sleep quality, latency, duration, efficiency, and disturbance, but not medication. However, the Spanish participants reported more stable PROs.

Women and men reported similar levels of sleep quality with equal medians and means (0.9 and 0.8). Question Q5A: Trouble sleeping: cannot get to sleep influenced the quality of sleep in women, as observed by a difference of over one unit from a maximum of 3 between means. Males provided more stable results with a standard deviation of 2.45 compared to 4.32 for the numeric score. At the extremity of inadequate sleep, the worst six levels of sleep quality correspond to women from both Spain and Hungary.

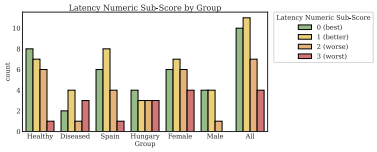
Table A9. Characteristics of PRO Sleep (PSQI).

ID	Health	Wave	Country	Gender	Age	Q1: Time gone to bed at night	Q2: Duration taken to fall asleep	Q3: Time gotten up in the morning	Q4: Duration of actual sleep	Q5A: Trouble sleeping: cannot get to sleep	Q5B: Trouble sleeping: wake up in the middle of the night	Q5C: Trouble sleeping: use the bathroom	Q5D: Trouble sleeping: cannot breathe comfortably	Q5E: Trouble sleeping: cough or snore loudly	Q5F: Trouble sleeping: too cold	Q5G: Trouble sleeping: too hot	Q5H: Trouble sleeping: bad dreams	Q5I: Trouble sleeping: pain	Q5J: Trouble sleeping for other reason(s)	Q6: Frequency of medicine to help you sleep	Q7: Trouble staying awake while driving, eating, or socializing	Q8: Problem with keeping up enthusiasm to get things done	Q9: Sleep quality overall	Subjective Quality Numeric Sub-Score	Latency Numeric Sub-Score	Duration Numeric Sub-Score	Efficiency Numeric Sub-Score	Disturbance Numeric Sub-Score	Medication Numeric Sub-Score	Daytime Dysfunction Numeric Sub-Score	Efficiency Numeric Sub-Score	**Numeric Score**	Categorical Score
535	Healthy	3	Hungary	Male	69	1410.0	10.0	510.0	480.0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0.9	1	1
643	Healthy	1	Spain	Female	67	1410.0	5.0	420.0	420.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.9	1	1
628	Healthy	1	Spain	Female	70	1440.0	10.0	510.0	480.0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0.9	1	1
648	Healthy	1	Spain	Female	72	1500.0	10.0	540.0	480.0	0	1	2	0	0	0	2	0	1	0	0	0	0	0	0	0	0	1	0	0	1.0	1	1
620	Healthy	2	Spain	Female	69	1435.0	0.0	480.0	480.0	0	1	0	0	0	0	0	1	1	1	0	0	0	1	1	0	0	0	1	0	0	1.0	2	1
575	Healthy	2	Hungary	Female	65	1395.0	5.0	420.0	420.0	0	1	3	0	0	0	0	3	1	0	0	1	0	0	0	0	1	0	1	0	1	0.9	3	1
634	Diseased	1	Spain	Male	72	1439.0	15.0	420.0	420.0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	1.0	3	1
569	Healthy	3	Hungary	Female	67	1385.0	10.0	405.0	420.0	2	2	3	0	2	1	0	1	0	0	0	0	0	1	1	1	1	0	1	0	0	0.9	4	1
133	Healthy	3	Hungary	Female	71	1315.0	30.0	375.0	480.0	2	0	1	0	2	0	0	1	2	0	0	0	0	1	1	2	0	0	1	0	0	1.0	4	1
132	Diseased	3	Hungary	Male	71	1330.0	10.0	370.0	420.0	0	2	0	0	0	1	1	2	0	0	1	1	1	1	0	1	0	1	0	1	0	1	0.9	4	1
800	Diseased	3	Spain	Female	65	1440.0	15.0	420.0	360.0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	0	0	0	0	0.9	4	1
643	Healthy	2	Spain	Female	67	1380.0	1.0	435.0	420.0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	1	1	0	0	0.8	4	1
790	Healthy	3	Spain	Male	66	1440.0	2.0	420.0	420.0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	0	1	0	1	0	1	1.0	4	1
803	Healthy	3	Spain	Female	67	1500.0	15.0	390.0	360.0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	1	2	0	1	0	1	1.0	5	0
638	Healthy	1	Spain	Female	71	1470.0	24.0	480.0	420.0	1	0	0	0	1	0	1	0	0	0	1	0	0	1	1	1	1	0	1	1	0	0.9	5	0
133	Healthy	1	Hungary	Female	71	1375.0	30.0	353.0	420.0	2	1	2	0	2	0	2	1	0	0	0	0	1	1	1	2	1	0	1	0	1	1.0	6	0
170	Healthy	3	Hungary	Male	70	1380.0	5.0	360.0	360.0	1	2	2	0	0	0	1	0	0	1	0	1	1	1	1	1	2	0	1	0	1	0.9	6	0
640	Healthy	1	Spain	Female	69	1440.0	20.0	480.0	420.0	2	2	2	0	2	0	3	0	1	0	0	0	0	1	1	2	1	0	2	0	0	0.9	6	0
795	Healthy	3	Spain	Female	72	1500.0	10.0	510.0	360.0	2	2	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	1	1	0	0	0.8	6	0
420	Healthy	2	Hungary	Female	71	1350.0	20.0	510.0	480.0	2	2	3	1	2	0	0	2	0	0	0	0	1	1	1	2	0	1	2	0	1	0.8	7	0
419	Healthy	2	Hungary	Male	95	1320.0	3.0	330.0	360.0	0	2	3	0	0	0	0	0	2	2	0	2	1	1	1	0	2	1	1	0	2	0.8	7	0
799	Diseased	3	Spain	Male	79	1440.0	30.0	600.0	360.0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	2	3	0	0	1	0.6	7	0
641	Diseased	2	Spain	Female	71	1380.0	5.0	394.0	360.0	0	1	2	0	0	0	1	0	0	0	3	0	0	1	1	0	2	1	1	3	0	0.8	8	0
624	Diseased	1	Spain	Female	72	1410.0	30.0	465.0	420.0	2	3	3	0	2	0	0	0	2	0	0	1	0	1	1	2	1	1	2	0	1	0.8	8	0
625	Diseased	1	Spain	Male	72	1380.0	15.0	360.0	360.0	0	3	1	0	3	0	0	0	3	0	0	0	0	3	3	1	2	0	2	0	0	0.9	8	0
700	Healthy	2	Spain	Male	67	1410.0	15.0	480.0	360.0	2	3	3	0	0	0	0	0	0	1	0	1	0	1	1	2	2	2	1	0	1	0.7	9	0
617	Healthy	2	Spain	Female	69	1416.0	30.0	517.0	360.0	3	3	2	1	1	1	0	2	1	0	1	0	0	1	1	2	2	2	2	1	0	0.7	10	0
215	Healthy	1	Hungary	Female	87	1380.0	30.0	420.0	250.0	1	1	3	1	0	0	1	1	1	0	1	0	1	2	2	1	3	3	1	1	1	0.5	12	0
630	Healthy	1	Spain	Female	74	1380.0	60.0	390.0	300.0	3	2	2	0	0	0	3	0	0	2	3	0	0	1	1	3	2	2	1	3	0	0.7	12	0
169	Diseased	3	Hungary	Female	69	1380.0	100.0	210.0	150.0	3	3	1	0	3	1	2	2	1	0	0	0	1	2	2	3	3	3	2	0	1	0.6	14	0
169	Diseased	1	Hungary	Female	69	1404.0	60.0	245.0	150.0	3	3	1	0	3	0	2	2	3	3	0	0	1	3	3	3	3	3	2	0	1	0.5	15	0
169	Diseased	2	Hungary	Female	69	1320.0	120.0	300.0	161.0	3	3	1	1	3	3	1	2	3	3	0	0	1	2	2	3	3	3	3	0	1	0.4	15	0
Median: Healthy						1410.0	10.0	427.5	420.0	1.0	1.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.8	5.0	0.0
Median: Diseased						1392.0	22.5	382.0	360.0	0.0	2.5	1.0	0.0	1.0	0.0	0.5	0.0	1.5	0.0	0.0	0.0	0.0	1.0	1.0	1.0	2.0	1.0	1.5	0.0	1.0	0.8	8.0	0.0
Median: Spain						1439.0	15.0	465.0	420.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.8	5.0	0.0
Median: Hungary						1380.0	20.0	370.0	420.0	2.0	2.0	2.0	0.0	2.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	1.0	0.8	6.0	0.0
Median: Female						1404.0	20.0	420.0	420.0	2.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.8	6.0	0.0
Median: Male						1410.0	10.0	420.0	360.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	1.0	1.0	2.0	0.0	1.0	0.0	1.0	0.8	6.0	0.0
Median: All						1407.0	15.0	420.0	420.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	1.0	0.0	0.0	0.8	6.0	0.0
Mean: Healthy						1410.5	15.6	442.5	406.8	1.0	1.2	1.4	0.1	0.5	0.0	0.6	0.5	0.5	0.3	0.2	0.3	0.2	0.7	0.7	1.0	1.1	0.5	1.0	0.2	0.4	0.8	5.2	0.4
Mean: Diseased						1392.3	40.0	378.4	316.1	1.1	1.8	0.9	0.1	1.4	0.4	0.7	0.7	1.4	0.6	0.3	0.4	0.4	1.5	1.5	1.5	2.0	1.4	1.3	0.3	0.6	0.7	8.6	0.3
Mean: Spain						1432.1	16.4	458.4	397.8	0.8	1.2	0.9	0.0	0.4	0.0	0.5	0.2	0.5	0.2	0.4	0.3	0.0	0.7	0.7	1.0	1.3	0.6	1.0	0.4	0.2	0.8	5.4	0.4
Mean: Hungary						1364.9	33.3	369.8	350.0	1.4	1.7	1.7	0.2	1.3	0.3	0.7	1.2	1.1	0.6	0.0	0.3	0.6	1.2	1.2	1.3	1.5	1.0	1.3	0.0	0.8	0.7	7.5	0.3
Mean: Female						1408.9	27.8	420.3	372.6	1.3	1.4	1.3	0.1	1.0	0.2	0.8	0.7	0.8	0.3	0.3	0.1	0.3	0.9	0.9	1.3	1.3	0.9	1.2	0.3	0.3	0.8	6.6	0.3
Mean: Male						1394.3	11.6	427.7	393.3	0.3	1.4	1.0	0.0	0.3	0.0	0.2	0.2	0.7	0.4	0.0	0.8	0.3	1.0	1.0	0.6	1.4	0.6	0.8	0.0	0.7	0.8	5.4	0.4
Mean: All						1404.8	23.2	422.4	378.4	1.0	1.4	1.2	0.1	0.8	0.1	0.6	0.6	0.8	0.4	0.2	0.3	0.3	0.9	0.9	1.1	1.4	0.8	1.1	0.2	0.5	0.8	6.3	0.4
SD: Healthy						51.1	13.8	60.2	61.3	1.0	0.9	1.2	0.3	0.8	0.2	0.9	0.8	0.7	0.6	0.6	0.5	0.4	0.5	0.5	0.9	0.8	0.8	0.4	0.6	0.5	0.1	3.2	0.4
SD: Diseased						40.9	38.2	106.3	109.2	1.3	1.3	0.9	0.3	1.4	0.9	0.7	0.9	1.2	1.2	0.9	0.6	0.4	0.9	0.9	1.1	0.7	1.3	1.0	0.9	0.4	0.1	4.3	0.4
SD: Spain						38.5	13.8	59.3	48.5	1.0	1.1	1.0	0.2	0.8	0.2	0.9	0.5	0.8	0.5	0.9	0.5	0.2	0.6	0.6	0.8	0.7	0.9	0.6	0.9	0.4	0.1	3.1	0.4
SD: Hungary						32.3	36.2	84.8	122.8	1.1	0.8	1.1	0.4	1.2	0.8	0.7	0.8	1.0	1.1	0.2	0.6	0.4	0.7	0.7	1.1	1.1	1.3	0.6	0.2	0.5	0.1	4.6	0.4
SD: Female						50.4	29.9	83.2	102.2	1.1	1.0	1.1	0.3	1.1	0.6	1.0	0.9	0.9	0.9	0.8	0.3	0.4	0.7	0.7	1.0	1.0	1.1	0.6	0.8	0.4	0.1	4.3	0.4
SD: Male						43.1	8.0	82.5	41.0	0.6	1.1	1.2	0.0	0.9	0.0	0.4	0.4	1.1	0.6	0.0	0.7	0.4	0.8	0.8	0.6	0.6	1.0	0.5	0.0	0.6	0.1	2.4	0.4
SD: All						48.9	26.7	83.1	89.8	1.1	1.0	1.1	0.3	1.1	0.5	0.9	0.8	1.0	0.8	0.7	0.5	0.4	0.7	0.7	1.0	0.9	1.1	0.6	0.7	0.5	0.1	3.9	0.4

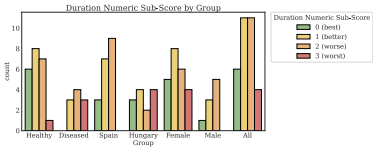
Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome).



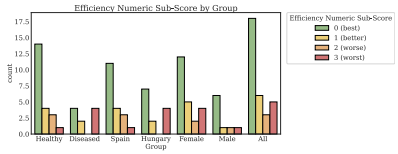
(a) Subjective Quality Numeric Sub-Score



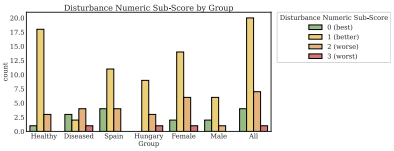
(b) Latency Numeric Sub-Score



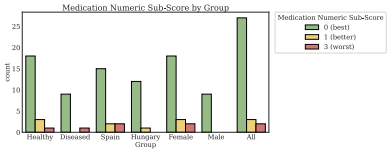
(c) Duration Numeric Sub-Score



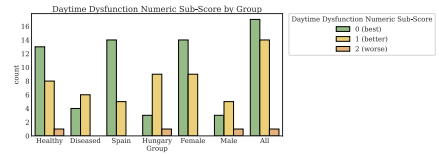
(d) Efficiency Numeric Sub-Score



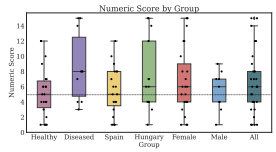
(e) Disturbance Numeric Sub-Score



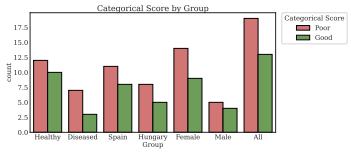
(f) Medication Numeric Sub-Score



(g) Daytime Dysfunction Numeric Sub-Score



(h) Numeric Score



(i) Categorical Score

Figure A9. Sub-scores and Scores for Sleep (PSQI). Dotted markings delimit levels of the categorical score.

Health-Related Quality of Life (EQ-5D-3L)

Participants provided 30 answers about their quality of life on the EQ-5D-3L scale [34]. The scale provides 3 severity levels for five facets of life quality, *no problem*, *some problems*, and *extreme problems* as well as a 0–100 numeric score for the health status on the day of the administration, as detailed in Appendix B.1.1. Half of the answers report a health score of 90 or above. Five answers reported a health score of 75 or below, and five answers reported a health score of 100. Table A10 shows the answers and Figure A10 illustrates the sub-scores and scores by participant group.

(a) Mobility Sub-Score

(b) Self-Care Sub-Score

(c) Usual Activities Sub-Score

(d) Pain and Discomfort Sub-Score

(e) Anxiety and Depression Sub-Score

(f) Health Score

Figure A10. Sub-scores and Scores for Health-Related Quality of Life (EQ-5D-3L).

The mean ± SD perceived health is at 84.96 ± 13.8 across all participants. The means ± SD for the five domains are as follows: 1.2 ± 0.4 for mobility, 1.0 ± 0.0 for self-care, 1.1 ± 0.3 for usual activities, 1.5 ± 0.6 for pain/discomfort, and 1.2 ± 0.4 for depression/anxiety. None of the participants self-reported quality of life issues due to self-care impediments.

The healthy and diseased participants report similar quality of life in the mobility, self-care, and usual activities. However, the participants with disease report worse pain/discomfort and depression/anxiety. Furthermore, the participants with disease report a mean health score of only 77.27 as compared to the 89.42 for the healthy. The participants with disease also self-report less stable answers, e.g., SD for the health score of 16.97 as compared to the SD of 8.95 of the healthy.

Participants from Spain self-reported a slightly improved health than those from Hungary. The participants from Spain reported a median health score of 90 compared to 85 for those from Hungary. However, the mean health scores are similar: 86.84 and 83.52, respectively. The participants from Hungary participants provided more stable health score, but more varied depression/anxiety responses than the participants from Spain.

Female participants report similar health as compared to male participants, with a median health score of 85 compared to 90, but a mean of 85.42 compared to 83.88. Women self-report experiencing slightly less mobility, usual activities, and depression/anxiety.

Table A10. Characteristics of PRO Health-Related Quality of Life (EQ-5D-3L).

ID	Health	Wave	Country	Gender	Age	Q1: Mobility	Q2: Self-Care	Q3: Usual Activities	Q4: Pain/Discomfort	Q5: Anxiety/Depression	Health Score
625	Diseased	1	Spain	Male	72	1	1	1	3	1	40
641	Diseased	1	Spain	Female	71	1	1	1	2	2	50
640	Healthy	1	Spain	Female	69	1	1	1	2	1	70
169	Diseased	2	Hungary	Female	69	1	1	2	2	2	75
630	Healthy	1	Spain	Female	74	2	1	2	1	2	75
169	Diseased	1	Hungary	Female	69	2	1	1	2	2	80
169	Diseased	3	Hungary	Female	69	1	1	1	1	2	80
420	Healthy	2	Hungary	Female	71	2	1	1	2	2	80
132	Diseased	1	Hungary	Male	71	1	1	1	1	1	80
641	Diseased	2	Spain	Female	71	1	1	1	2	1	80
624	Diseased	1	Spain	Female	72	2	1	1	2	1	80
648	Healthy	1	Spain	Female	72	2	1	2	2	1	80
796	Healthy	3	Spain	Male	74	1	1	1	1	1	80
575	Healthy	2	Hungary	Female	65	1	1	1	2	1	85
170	Healthy	3	Hungary	Male	70	1	1	1	2	2	85
132	Diseased	3	Hungary	Male	71	1	1	1	2	1	90
212	Healthy	1	Hungary	Male	72	1	1	2	1	1	90
643	Healthy	2	Spain	Female	67	1	1	1	1	1	90
617	Healthy	2	Spain	Female	69	1	1	1	1	1	90
795	Healthy	3	Spain	Female	72	1	1	1	1	1	90
569	Healthy	3	Hungary	Female	67	1	1	1	2	1	95
133	Healthy	1	Hungary	Female	71	1	1	1	1	1	95
419	Healthy	2	Hungary	Male	95	1	1	1	1	1	95
799	Diseased	3	Spain	Male	79	1	1	1	1	1	95
133	Healthy	3	Hungary	Female	71	1	1	1	1	1	99
800	Diseased	3	Spain	Female	65	1	1	1	1	1	100
643	Healthy	1	Spain	Female	67	1	1	1	1	1	100
628	Healthy	1	Spain	Female	70	1	1	1	1	1	100
638	Healthy	1	Spain	Female	71	1	1	1	1	1	100
790	Healthy	3	Spain	Male	66	1	1	1	1	1	100
Median: Healthy						1.0	1.0	1.0	1.0	1.0	90.0
Median: Diseased						1.0	1.0	1.0	2.0	1.0	80.0
Median: Spain						1.0	1.0	1.0	1.0	1.0	90.0
Median: Hungary						1.0	1.0	1.0	2.0	1.0	85.0
Median: Female						1.0	1.0	1.0	1.0	1.0	85.0
Median: Male						1.0	1.0	1.0	1.0	1.0	90.0
Median: All						1.0	1.0	1.0	1.0	1.0	87.5
Mean: Healthy						1.1	1.0	1.1	1.3	1.1	89.4
Mean: Diseased						1.1	1.0	1.0	1.7	1.3	77.2
Mean: Spain						1.1	1.0	1.1	1.4	1.1	83.5
Mean: Hungary						1.1	1.0	1.1	1.5	1.3	86.8
Mean: Female						1.2	1.0	1.1	1.4	1.2	85.4
Mean: Male						1.0	1.0	1.1	1.4	1.1	83.8
Mean: All						1.1	1.0	1.1	1.4	1.2	84.9
SD: Healthy						0.3	0.0	0.3	0.4	0.3	8.9
SD: Diseased						0.3	0.0	0.2	0.6	0.4	16.9
SD: Spain						0.3	0.0	0.3	0.5	0.3	17.0
SD: Hungary						0.3	0.0	0.3	0.4	0.4	7.3
SD: Female						0.4	0.0	0.3	0.4	0.4	12.2
SD: Male						0.0	0.0	0.3	0.6	0.3	16.7
SD: All						0.3	0.0	0.3	0.5	0.4	13.8

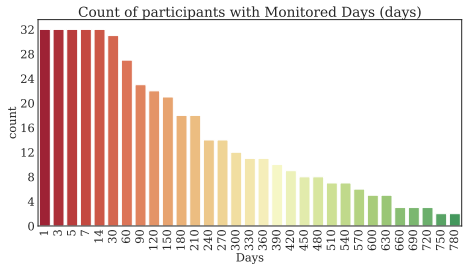
Color coding: from orange (worse outcome relative to others) to yellow to reen (better outcome).

Appendix C.1.2. Technology-Reported Outcomes (Fitbit)

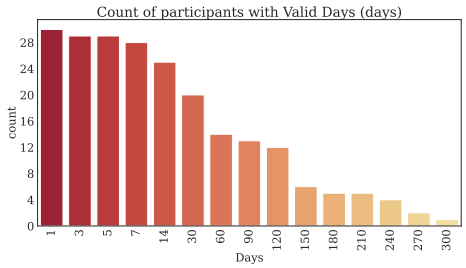
We overview the TechROs by first assessing the data quality. Table A11 depicts the total compliance (as the number of days including TechROs) as well as the intraday compliance (as the number of valid days). Figure A11 depicts participant compliance in days (all monitored and valid) for each participant group. Figure A12 illustrates participant compliance by outcome. Figures A13–A15 show participant compliance by health, country, and gender groups, respectively.

While participants wore the devices for a median (mean) of 224 (295) days, Fitbit reported TechROs for different durations. Energy expenditure, steps, and heart rate appeared in the majority of days,

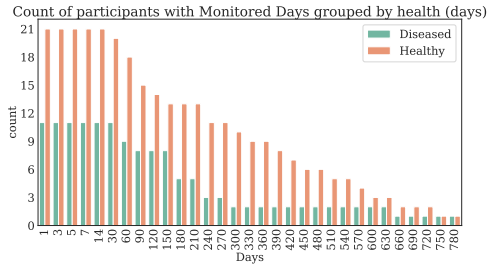
with their medians (means ± SD) at 224, 204, and 128 (295 ± 238, 276 ± 236, and 230 ± 214) days. The sedentary, light, fair, and vigorous physical activity durations appeared in decreasing durations, with medians (means ± SD) at 136, 136, 91, and 79 days (219 ± 203, 219 ± 202, 165 ± 171, and 160 ± 168 days). Sleep monitoring recorded a median (mean ± SD) of 130 (198 ± 194) days. Cumulative TechROs such as sedentary+light recorded durations corresponding to at most the minimum of their constituents.



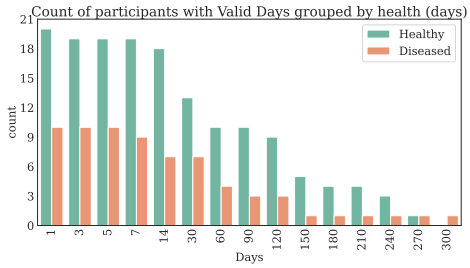
(a) All Days, All Participants



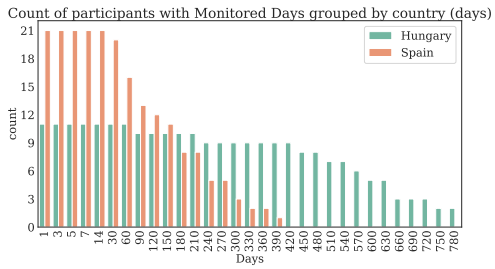
(b) Valid Days, All Participants



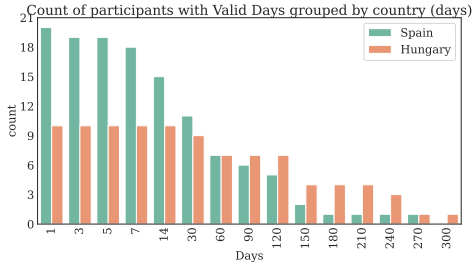
(c) All Days, by Health



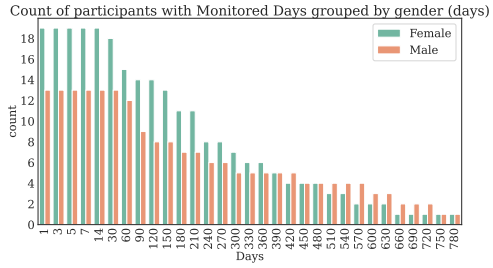
(d) Valid Days, by Health



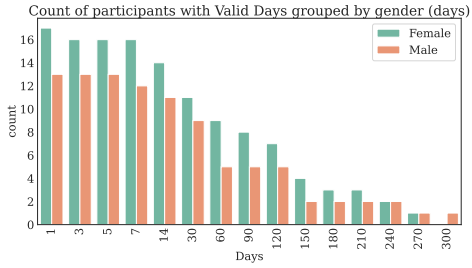
(e) All Days, by Country



(f) Valid Days, by Country

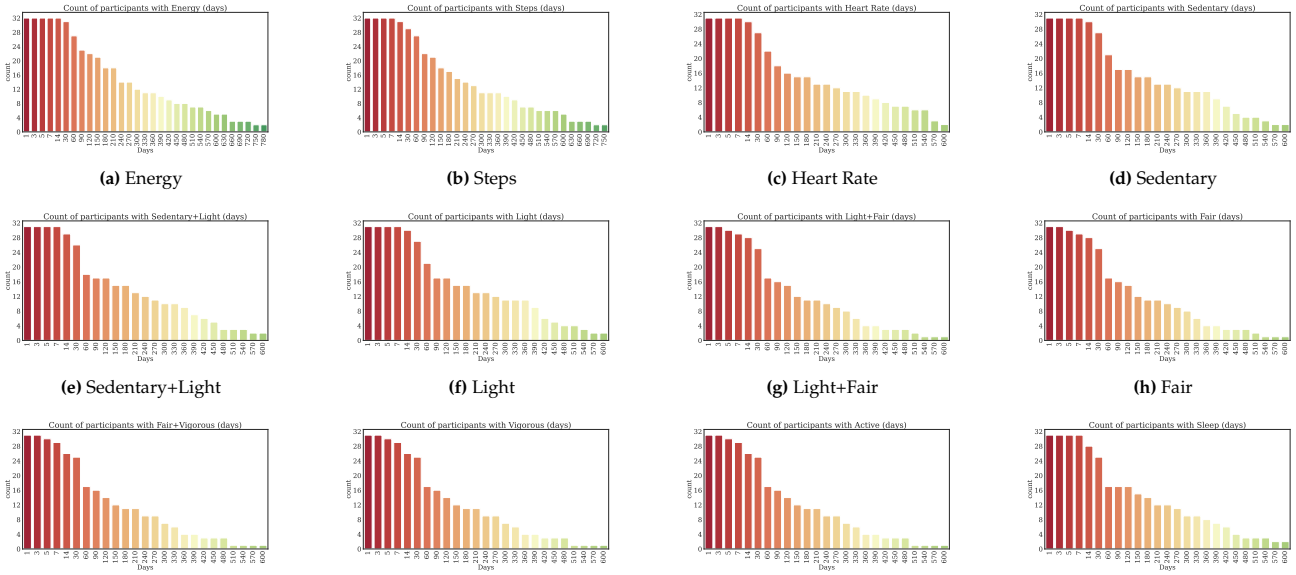


(g) All Days, by Gender



(h) Valid Days, by Gender

Figure A11. Count of seniors with at least the given valid days of Fitbit (TechRO) by group.

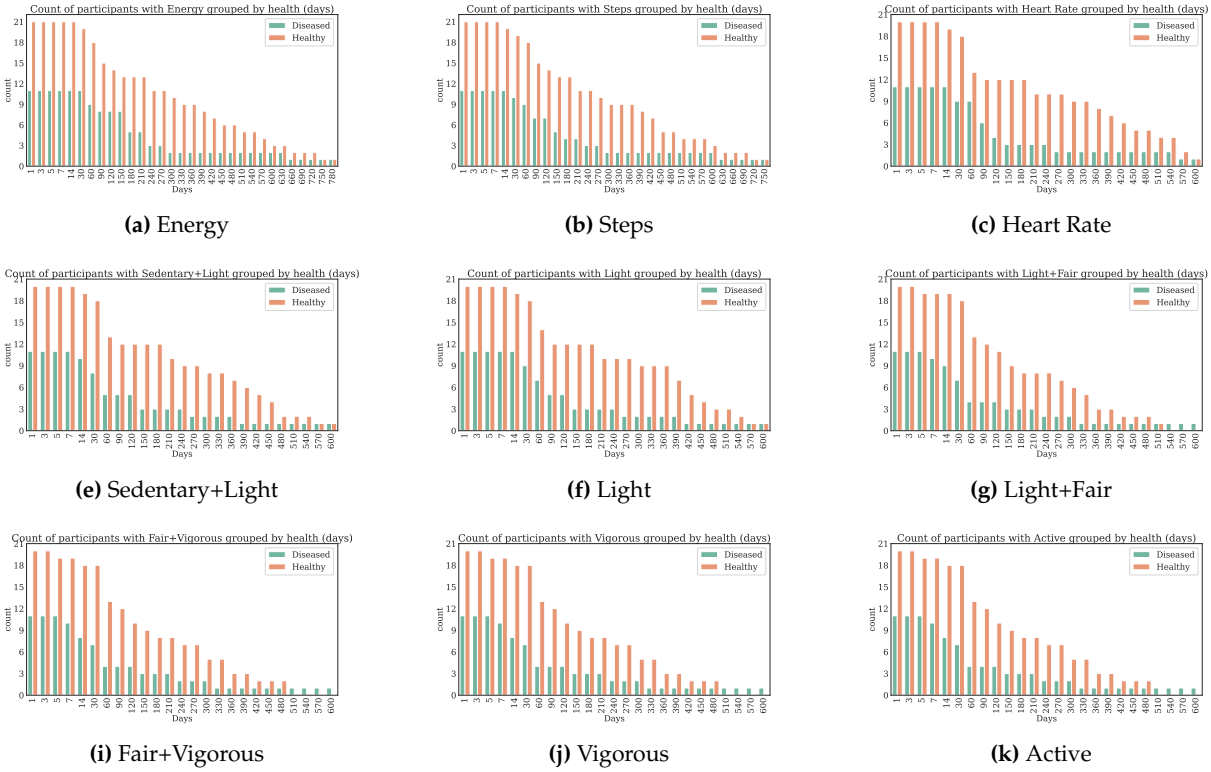


(a) Energy **(b)** Steps **(c)** Heart Rate **(d)** Sedentary

(e) Sedentary+Light **(f)** Light **(g)** Light+Fair **(h)** Fair

(i) Fair+Vigorous **(j)** Vigorous **(k)** Active **(l)** Sleep

Figure A12. Count of seniors with at least the given valid days of Fitbit (TechRO) by outcome.

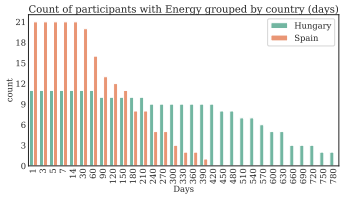


(a) Energy **(b)** Steps **(c)** Heart Rate

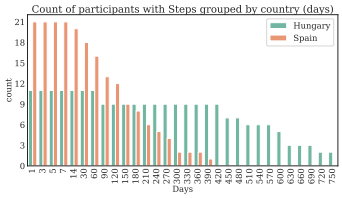
(e) Sedentary+Light **(f)** Light **(g)** Light+Fair

(i) Fair+Vigorous **(j)** Vigorous **(k)** Active

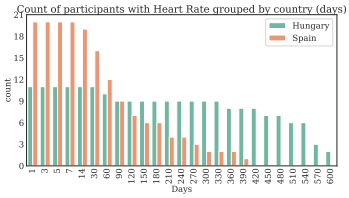
Figure A13. Count of seniors with at least the given valid days of Fitbit (TechRO) by outcome and health group.



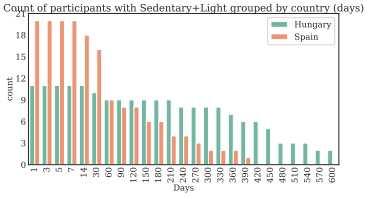
(a) Energy



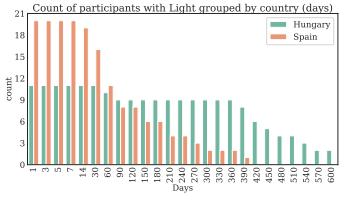
(b) Steps



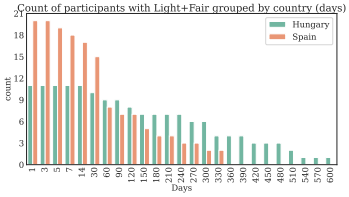
(c) Heart Rate



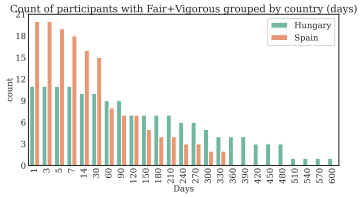
(e) Sedentary+Light



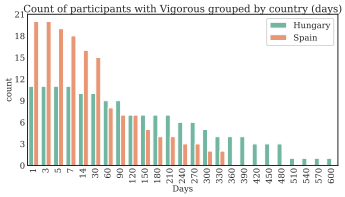
(f) Light



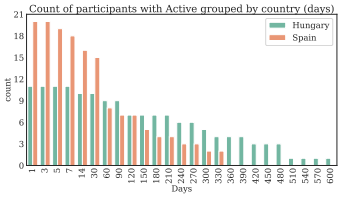
(g) Light+Fair



(i) Fair+Vigorous

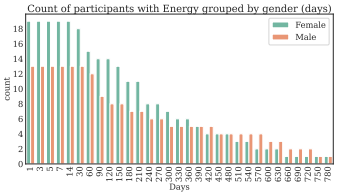


(j) Vigorous

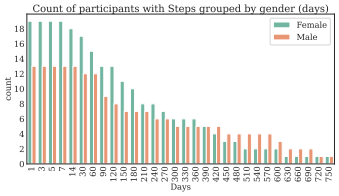


(k) Active

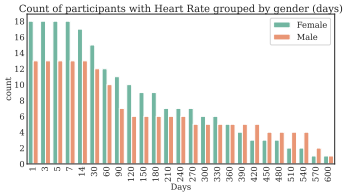
Figure A14. Count of seniors with at least the given valid days of Fitbit (TechRO) by outcome and country group.



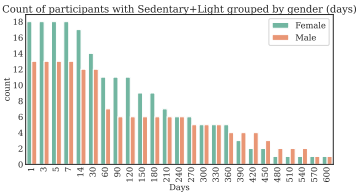
(a) Energy



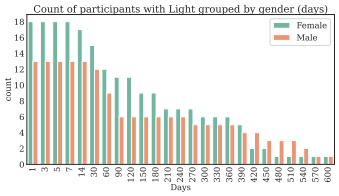
(b) Steps



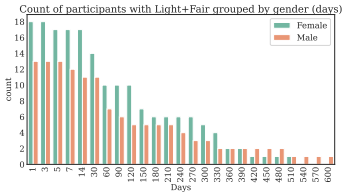
(c) Heart Rate



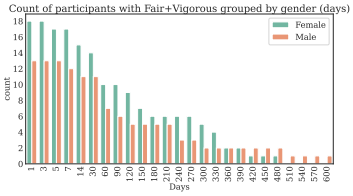
(e) Sedentary+Light



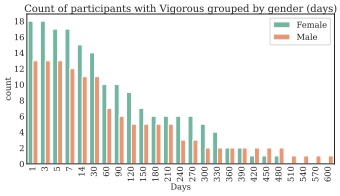
(f) Light



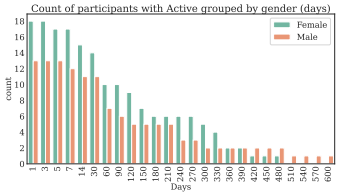
(g) Light+Fair



(i) Fair+Vigorous



(j) Vigorous



(k) Active

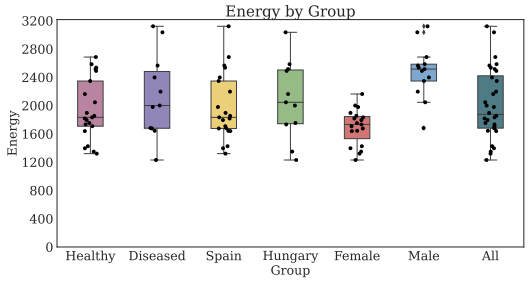
Figure A15. Count of seniors with at least the given valid days of Fitbit (TechRO) by outcome and gender group.

Table A11. Days of Fitbit (TechRO) data for seniors with at least one PRO ($N = 32$ participants).

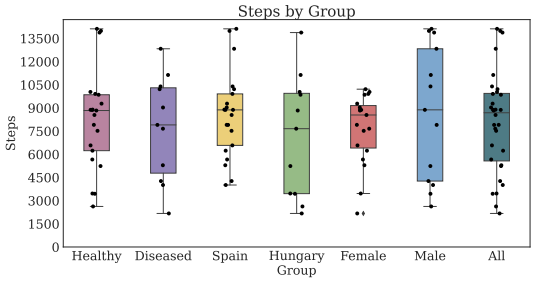
PID	Health	Country	Gender	Age	Monitored Days	Energy	Steps	Heart Rate	Sedentary	Sedentary+Light	Light	Light+Fair	Fair	Fair+Very	Very	Active	Sleep	Valid Days
502	Diseased	Hungary	Female	63	231	231	80	73	50	22	49	15	15	9	9	9	8	0
649	Healthy	Spain	Female	72	135	135	125	0	0	0	0	0	0	0	0	0	0	0
630	Healthy	Spain	Female	74	23	23	13	7	7	7	7	4	4	4	4	4	7	1
799	Diseased	Spain	Male	79	34	34	20	15	16	13	15	6	6	6	6	6	13	5
648	Healthy	Spain	Female	72	35	35	24	21	22	19	22	14	14	12	12	12	18	11
791	Diseased	Spain	Male	72	74	74	67	64	56	54	56	12	12	12	12	12	54	11
800	Diseased	Spain	Female	65	43	43	33	22	22	20	22	17	17	17	17	17	19	12
798	Healthy	Spain	Female	67	47	47	39	35	36	35	36	31	31	31	31	31	34	20
796	Healthy	Spain	Male	74	77	77	63	46	49	48	49	48	48	48	48	48	25	20
575	Healthy	Hungary	Female	65	69	69	61	59	60	58	60	41	41	41	41	41	59	23
795	Healthy	Spain	Female	72	274	274	261	40	38	38	38	32	32	31	31	31	35	26
790	Healthy	Spain	Male	66	79	79	70	67	67	65	67	63	63	63	63	63	30	28
624	Diseased	Spain	Female	72	153	153	143	138	139	130	139	50	50	47	47	47	131	30
420	Healthy	Hungary	Female	71	552	552	446	344	420	233	417	138	138	113	114	113	173	34
644	Diseased	Spain	Male	70	169	169	142	90	61	58	61	53	53	53	53	53	54	37
576	Healthy	Hungary	Male	60	439	439	430	430	430	420	430	119	119	94	95	94	429	49
634	Diseased	Spain	Male	72	237	237	230	75	60	57	60	55	55	55	55	55	56	50
793	Healthy	Spain	Male	68	119	119	107	52	52	52	52	52	52	52	52	52	51	51
641	Diseased	Spain	Female	71	167	167	156	118	134	132	134	122	122	120	121	120	129	63
643	Healthy	Spain	Female	67	217	217	201	186	186	186	186	171	171	159	159	159	185	90
638	Healthy	Spain	Female	71	211	211	208	208	208	207	208	147	147	129	130	129	206	127
212	Healthy	Hungary	Male	72	733	733	698	580	538	465	538	244	244	230	233	230	439	136
170	Healthy	Hungary	Male	70	785	785	777	551	369	353	363	303	303	298	299	298	347	140
169	Diseased	Hungary	Female	69	794	794	778	561	398	360	398	312	312	302	303	302	293	141
625	Diseased	Spain	Male	72	288	288	276	254	254	250	254	221	221	217	217	217	250	141
628	Healthy	Spain	Female	70	303	303	290	286	289	288	289	278	278	273	273	273	276	146
617	Healthy	Spain	Female	69	402	402	395	391	392	392	392	355	355	342	344	342	392	170
569	Healthy	Hungary	Female	67	501	501	498	483	479	475	479	417	417	415	415	415	476	215
133	Healthy	Hungary	Female	71	632	632	623	622	623	622	623	521	521	486	487	486	621	242
419	Healthy	Hungary	Male	95	599	599	594	561	567	563	566	502	502	493	494	493	553	245
640	Healthy	Spain	Female	69	385	385	380	376	377	375	377	346	346	344	345	344	371	295
132	Diseased	Hungary	Male	71	639	639	623	618	622	619	622	613	613	613	613	613	617	301
Median: Healthy					274.0	274.0	261.0	208.0	208.0	207.0	208.0	138.0	138.0	113.0	114.0	113.0	185.0	51.0
Median: Diseased					169.0	169.0	143.0	90.0	61.0	58.0	61.0	53.0	53.0	53.0	53.0	53.0	56.0	37.0
Median: Spain					153.0	153.0	142.0	67.0	60.0	57.0	60.0	52.0	52.0	52.0	52.0	52.0	54.0	30.0
Median: Hungary					599.0	599.0	594.0	551.0	430.0	420.0	430.0	303.0	303.0	298.0	299.0	298.0	429.0	140.0
Median: Female					217.0	217.0	201.0	138.0	139.0	132.0	139.0	122.0	122.0	113.0	114.0	113.0	131.0	34.0
Median: Male					237.0	237.0	230.0	90.0	67.0	65.0	67.0	63.0	63.0	63.0	63.0	63.0	56.0	50.0
Median: All					224.0	224.0	204.5	128.0	136.5	131.0	136.5	91.0	91.0	78.5	79.0	78.5	130.0	49.5
Mean: Healthy					315.0	315.0	300.1	254.5	248.0	233.3	247.5	182.1	182.1	174.1	174.7	174.1	225.0	98.5
Mean: Diseased					257.1	257.1	231.6	184.3	164.7	155.9	164.5	134.1	134.1	131.9	132.0	131.9	147.6	71.9
Mean: Spain					165.3	165.3	154.4	118.6	117.3	115.5	117.3	98.9	98.9	95.9	96.1	95.9	111.2	63.5
Mean: Hungary					543.0	543.0	509.8	443.8	414.1	380.9	413.1	293.1	293.1	281.2	282.0	281.2	365.0	138.7
Mean: Female					272.3	272.3	250.2	208.9	204.2	189.4	204.0	158.4	158.4	151.3	151.7	151.3	180.6	86.6
Mean: Male					328.6	328.6	315.1	261.7	241.6	232.0	241.0	176.2	176.2	171.8	172.3	171.8	224.4	93.3
Mean: All					295.1	295.1	276.5	230.4	219.4	206.7	219.0	165.6	165.6	159.6	160.0	159.6	198.4	89.3
SD: Healthy					238.7	238.7	234.0	216.8	207.8	198.5	207.5	164.9	164.9	160.6	161.0	160.6	199.4	89.8
SD: Diseased					231.9	231.9	235.4	200.9	181.4	179.0	181.6	177.8	177.8	177.2	177.3	177.2	173.8	86.6
SD: Spain					113.4	113.4	114.1	116.9	118.2	118.2	118.2	109.5	109.5	106.7	107.1	106.7	118.9	72.5
SD: Hungary					215.8	215.8	235.1	194.6	188.3	195.7	188.6	193.3	193.3	193.8	193.8	193.8	202.2	98.0
SD: Female					215.0	215.0	212.8	195.0	186.9	179.3	186.8	158.9	158.9	153.8	154.2	153.8	176.5	89.2
SD: Male					264.5	264.5	263.1	235.7	222.7	214.5	222.4	186.7	186.7	185.5	185.7	185.5	215.6	90.0
SD: All					238.0	238.0	236.7	214.1	203.0	195.5	202.8	171.0	171.0	167.7	168.0	167.7	194.5	89.6

Color coding: from orange (fewer days relative to others) to yellow to green (more days).

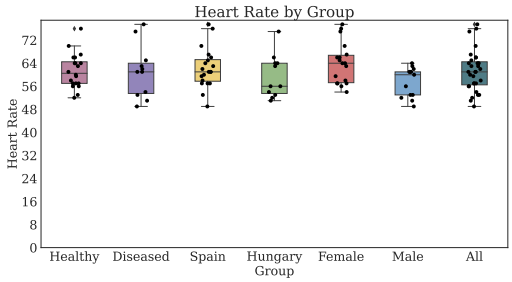
Concerning total compliance, Fitbit devices were worn by the participants in 295 ± 238 days on average and 50% of participants wore the Fitbit devices in at least 224 days. Healthy participants wore the devices on average 58 days more than participants with disease. Hungarian participants were also significantly more compliant in wearing the devices, by achieving mean 543 (446 more) days with monitored data. From the top 10 compliant, six were Hungarian. Most days were recorded by three Hungarians, and most valid days were recorded by one Hungarian. Men wore the devices for only slightly more extended periods than women.



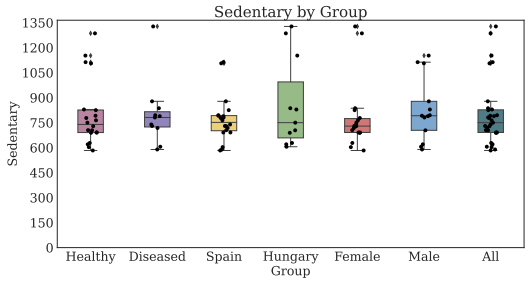
(a) Energy (kcal.)



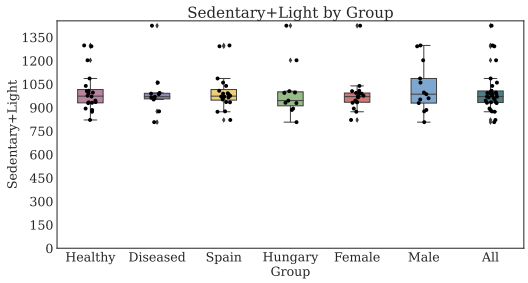
(b) Steps (count)



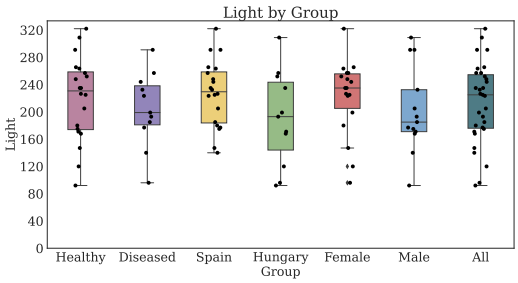
(c) Heart Rate (bpm.)



(d) Sedentary Duration (min.)



(e) Sedentary+Light Duration (min.)



(f) Light Duration (min.)

Figure A16. Median values of TechROs (Fitbit) across the entire monitoring period: energy, steps, heart rate, sedentary duration, sedentary+light duration, light duration (1 of 2).

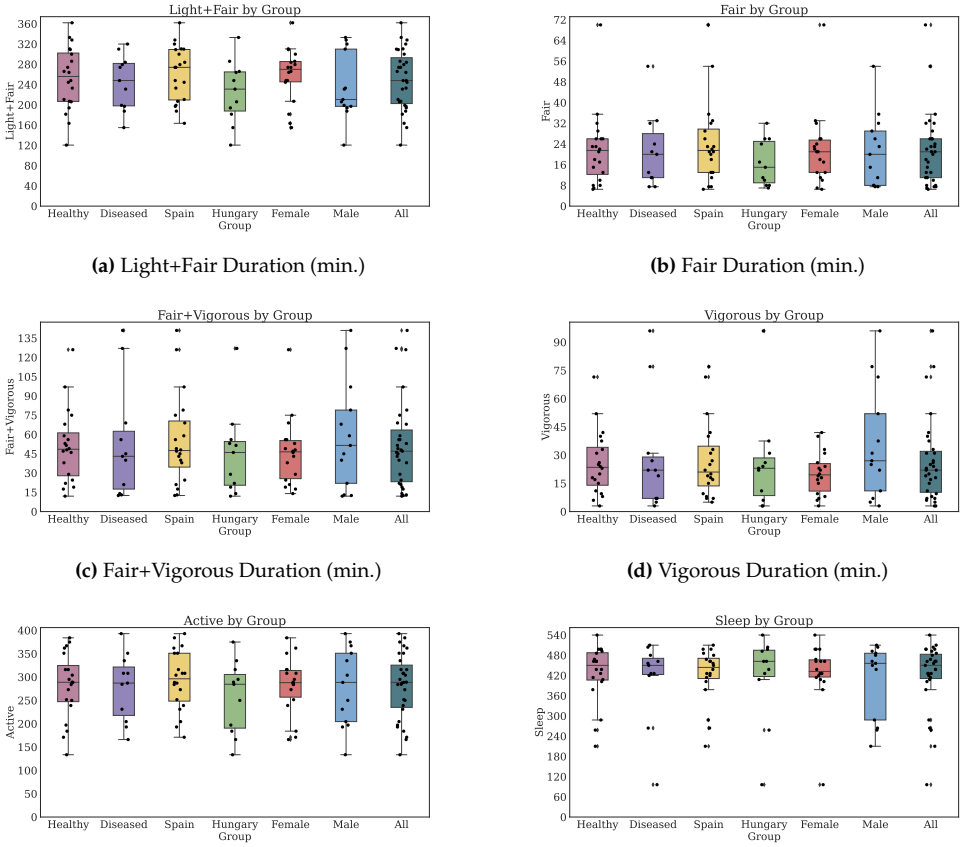
Regarding intraday compliance, participants wore the devices for more than 23 h for a mean ± SD of 89 ± 89 days while 50% of them wore the devices for at least 49 valid days of 21 h. One third had less than 30 valid days, half had less than 60 days, one person had 90 days, and one third had more than 120 days. The participants with disease were more compliant intraday than the healthy participants, keeping 37 valid days as compared to only 51 by the healthy participants, having a relative ratio to the total days of 4. Participants from Hungary were also more compliant intraday, achieving 140 valid days compared to 30 valid days and 13 ratio to total.

We overview the dataset by depicting in Table A12 the medians of the TechRO variables obtained from the participants' days over the entire period of monitoring and summary statistics by participant group. The following paragraphs describe each TechRO in depth. Figures 8 and 9 depict the median values for each group across the entire monitoring period.

Table A12. Median values of TechROs (Fitbit) across the entire monitoring period ($N = 32$ participants).

PID	Health	Country	Gender	Age	Energy (kcal.)	Steps (count)	Heart Rate (bpm)	Sedentary (min.)	Sedentary+Light (min.)	Light (min.)	Light+Fair (min.)	Fair (min.)	Fair+Very (min.)	Very (min.)	Active (min.)	Sleep (h:min.)
575	Healthy	Hungary	Female	65	1733.0	8835.0	64.0	750.0	1005.5	252.0	266.0	7.0	29.0	23.0	289.0	6:48
569	Healthy	Hungary	Female	67	1753.0	10,038.5	56.0	689.0	931.0	235.0	264.0	17.0	46.0	26.0	295.0	7:42
420	Healthy	Hungary	Female	71	1349.0	3462.0	66.0	1286.0	945.0	120.0	181.5	10.0	19.0	6.0	184.0	9:00
133	Healthy	Hungary	Female	71	2163.0	9856.0	64.0	628.0	894.0	257.0	286.0	26.0	53.0	24.0	316.0	8:18
576	Healthy	Hungary	Male	60	2516.0	2624.5	63.0	829.0	996.5	171.0	194.0	8.0	12.0	3.0	197.0	7:18
170	Healthy	Hungary	Male	70	2585.0	13,882.0	53.0	620.0	929.0	309.0	333.0	15.0	51.5	31.0	375.0	7:42
212	Healthy	Hungary	Male	72	2046.0	3445.0	56.0	1152.5	1203.0	92.0	120.5	8.0	22.0	11.0	133.5	4:18
419	Healthy	Hungary	Male	95	2490.0	5239.0	52.0	704.0	885.0	168.0	206.0	26.0	68.0	37.5	250.0	8:12
643	Healthy	Spain	Female	67	1795.0	9281.0	57.0	603.0	935.5	322.0	362.0	32.0	49.0	15.0	384.0	7:42
798	Healthy	Spain	Female	67	1817.0	9911.0	76.0	691.0	971.0	263.5	309.0	23.0	75.0	42.0	351.0	6:42
640	Healthy	Spain	Female	69	1708.0	8892.5	59.5	705.0	934.0	225.0	248.0	18.0	38.0	20.0	273.0	7:48
617	Healthy	Spain	Female	69	1639.0	8545.0	70.0	691.0	873.5	180.0	207.0	22.0	56.0	33.0	239.0	8:18
628	Healthy	Spain	Female	70	1833.0	8876.0	57.0	583.0	821.0	235.0	310.5	70.0	126.0	40.0	362.0	8:18
638	Healthy	Spain	Female	71	1896.0	7907.5	67.0	728.5	976.0	248.0	274.0	21.0	47.0	18.0	284.0	7:06
648	Healthy	Spain	Female	72	1425.0	6235.0	66.0	778.0	992.0	226.5	244.5	13.0	24.5	8.0	251.5	7:18
649	Healthy	Spain	Female	72	1854.0	7520.0										
795	Healthy	Spain	Female	72	1396.0	5664.0	58.0	764.0	1039.0	265.5	300.0	26.0	48.0	17.0	316.0	6:18
630	Healthy	Spain	Female	74	1320.0	6577.0	57.0	825.0	1008.0	147.0	163.5	6.5	17.5	9.5	171.0	6:54
790	Healthy	Spain	Male	66	2686.0	14123.5	60.0	1106.0	1298.0	205.0	233.0	23.0	79.0	52.0	304.0	3:30
793	Healthy	Spain	Male	68	2536.0	8879.0	64.0	791.5	1086.5	291.0	328.0	35.5	59.0	25.0	367.0	4:48
796	Healthy	Spain	Male	74	2347.0	13989.0	61.0	1113.0	1292.5	175.0	210.5	29.0	97.0	71.5	288.5	8:06
502	Diseased	Hungary	Female	63	1230.0	2171.0	75.0	1327.5	1424.5	96.0	155.0	11.0	14.0	3.0	166.0	1:36
169	Diseased	Hungary	Female	69	2000.5	7659.0	54.0	836.5	994.0	199.0	248.0	24.0	56.0	22.0	284.5	7:06
132	Diseased	Hungary	Male	71	3036.0	11136.0	51.0	605.5	807.0	193.0	231.0	32.0	127.0	96.0	335.0	8:24
800	Diseased	Spain	Female	65	1643.0	9030.0	77.5	739.0	989.5	244.0	284.0	21.0	43.0	19.0	308.0	7:00
641	Diseased	Spain	Female	71	1676.0	10216.0	65.0	718.0	965.5	223.5	274.0	33.0	69.0	31.0	308.0	7:06
624	Diseased	Spain	Female	72	1979.0	5292.0	63.0	730.0	970.0	257.0	279.5	13.0	21.0	7.0	287.0	7:42
644	Diseased	Spain	Male	70	2566.0	7903.5	61.0	781.0	952.0	177.0	197.0	11.0	40.0	27.0	231.0	7:30
625	Diseased	Spain	Male	72	2197.0	10394.5	53.0	589.0	876.0	291.0	320.0	20.0	45.0	22.0	351.0	8:30
634	Diseased	Spain	Male	72	3121.0	12832.5	61.0	794.5	1060.0	232.5	310.0	54.0	141.0	77.0	393.0	4:24
791	Diseased	Spain	Male	72	2397.5	4012.0	62.0	789.0	986.0	185.0	199.0	7.5	12.5	5.0	204.5	7:36
799	Diseased	Spain	Male	79	1682.0	4268.0	49.0	878.0	960.0	140.0	187.5	7.5	13.0	7.0	193.0	8:00
Median: Healthy					1833.0	8835.0	60.5	739.2	973.5	230.7	256.0	21.5	48.5	23.5	288.7	7:30
Median: Diseased					2000.5	7903.5	61.0	781.0	970.0	199.0	248.0	20.0	43.0	22.0	287.0	7:24
Median: Spain					1833.0	8876.0	61.0	751.5	973.5	229.5	274.0	21.5	47.5	21.0	296.2	7:24
Median: Hungary					2046.0	7659.0	56.0	750.0	945.0	193.0	231.0	15.0	46.0	23.0	284.5	7:36
Median: Female					1733.0	8545.0	64.0	729.2	970.5	235.0	270.0	21.0	46.5	19.5	288.0	7:12
Median: Male					2516.0	8879.0	60.0	791.5	986.0	185.0	210.5	20.0	51.5	27.0	288.5	7:30
Median: All					1875.0	8690.0	61.0	750.0	971.0	225.0	248.0	21.0	47.0	22.0	288.5	7:24
Mean: Healthy					1947.0	8275.3	61.3	801.8	1000.8	219.3	252.0	21.8	50.8	25.6	281.5	7:06
Mean: Diseased					2138.9	7719.5	61.0	798.9	998.5	203.4	244.0	21.2	52.8	28.7	278.2	6:48
Mean: Spain					1976.8	8588.0	62.2	769.8	999.3	226.6	262.0	24.3	55.0	27.3	293.3	7:00
Mean: Hungary					2081.9	7122.5	59.4	857.0	1001.3	190.1	225.9	16.7	45.2	25.6	256.8	6:54
Mean: Female					1695.2	7682.5	64.0	781.8	981.6	222.0	258.6	21.8	46.1	20.1	281.6	7:06
Mean: Male					2477.3	8671.4	57.3	827.1	1025.5	202.2	236.1	21.2	59.0	35.7	278.6	6:42
Mean: All					2012.9	8084.2	61.2	800.8	1000.0	213.7	249.2	21.6	51.5	26.7	280.3	7:00
SD: Healthy					423.2	3171.9	5.8	195.5	126.4	59.8	61.2	13.8	27.7	16.5	69.1	1:24
SD: Diseased					569.0	3237.4	8.7	186.8	148.5	52.3	51.7	13.4	42.2	28.9	68.1	1:54
SD: Spain					462.0	2735.1	6.7	135.3	115.4	47.6	53.6	15.2	34.1	19.9	62.4	1:18
SD: Hungary					524.4	3768.4	7.1	257.4	164.0	66.6	58.9	8.4	31.6	24.8	73.3	2:00
SD: Female					246.3	2204.8	6.9	196.6	119.0	54.1	52.0	13.9	25.8	11.0	59.4	1:30
SD: Male					363.2	4196.1	4.9	183.5	150.1	60.7	63.4	13.4	40.8	28.7	79.8	1:42
SD: All					486.9	3205.5	7.0	192.5	134.7	57.8	58.2	13.7	33.6	21.8	68.8	1:36

Color coding: from orange (worse outcome relative to others) to yellow to green (better outcome). Participant 649 only provided energy and steps.



(a) Light+Fair Duration (min.)

(b) Fair Duration (min.)

(c) Fair+Vigorous Duration (min.)

(d) Vigorous Duration (min.)

(e) Active Duration (min.)

(f) Sleep Duration (min.)

Figure A17. Median values of TechROs (Fitbit) across the entire monitoring period: light+fair duration, fair duration, fair+vigorous duration, vigorous duration, active duration, sleep duration (2 of 2).

Energy Expenditure (Raw Family)

For the energy expenditure Fitbit behavioural marker, participants spent a mean ± SD energy of 2013 ± 487 kcal. 50% participants spent 1896 kcal. or more per day. Table A12 illustrates these results.

Participants with disease consumed 100–200 kcal. more than healthy participants per day, with medians (means) of 2000 and 1825 (2139 and 1951). We observed a similar difference between the participants from Hungary and Spain (difference of means 213 kcal). Men consumed more calories than women, with respective medians (means) of 2516 and 1720 (2477 and 1686), but also with higher variation, with male SD 363 kcal. vs. female 250 kcal.

Steps (Raw Family)

For the steps Fitbit behavioural marker, participants were active: they performed a median (mean ± SD) of 8690 (8084 ± 3205) measured steps per day. Table A12 illustrates these results.

Healthy participants performed on average 556 more steps than participants with disease, and with a median difference of 932 steps. Healthy and diseased participants had comparable variabilities in the step counts. Participants from Spain performed on average 1217 more steps than participants from Hungary and the devices measured more consistency. Men performed 1992 more steps on average than women. However, the 50% step counts are similar, partly due to four males who performed more than 12.000 median steps per day.

Heart Rate (Raw Family)

For the heart rate behavioural marker measured by Fitbit, the median and (mean ± SD) were 61 (61 ± 7) beats per minute. Table A12 illustrates these results.

Both healthy and diseased participants reported similar heart rate means and medians. Devices owned by participants with disease reported higher variability between daily measures than healthy participants with 8.77 bpm. and 5.81 bpm., respectively. Hungarian participant devices reported a lower median at 56 compared to 61 bpm. On average, men had 3 bpm. less than women.

Sedentary Duration (Processed Family)

For the behavioural marker of sedentary duration, the participants recorded 801 ± 192 mean minutes per day. Table A12 illustrates these results.

Participants with disease report more sedentary time than healthy participants, with means of 781 and 739 min, respectively. Participants from Hungary report 88 min more sedentary duration on average with 857 compared to 769; however, they report similar medians. Men also report 242 min. more sedentary time than women, with medians 971 and 729 min, respectively.

Light Intensity Physical Activity Duration (Processed Family)

For the duration of physical activity at a light intensity as reported by Fitbit, all participants spend on average 213 ± 57 min per day. Table A12 illustrates these results.

Healthy participants report approximately 20 min more per day with a median (mean) of 230 (219) compared to 199 (203). Participants from Spain also report 30 min more with 229 median min for Spain compared to 193 median min for Hungary. Females are more active in the light intensity spectrum by 20 min than males.

Fair Intensity Physical Activity Duration (Processed Family)

For the duration of physical activity at a fair intensity as reported by Fitbit, all participants spend on average 21 ± 13 min per day. Table A12 illustrates these results.

Regardless of their grouping criteria of health status, country, or gender, participants consistently report means and medians in the 16–22 min for the fair intensity physical activity.

Vigorous Intensity Physical Activity Duration (Processed Family)

For the duration of physical activity at a vigorous intensity as reported by Fitbit, all participants spend on average 26 ± 21 min. per day. Table A12 illustrates these results.

Regardless of their grouping criteria of health status or country, participants consistently report means and medians in the 19–28 min for the vigorous-intensity physical activity. Men may perform vigorous physical activity for 10–15 min more than women, as observed in their respective medians (means) of 27 (35) and 19 (20), but also with more variability as their standard deviation is 28 compared to 11.

Sleep Duration (Processed Family)

For the sleep duration, participants sleep on average 7 ± 1.6 h and 50% of the participants sleep 7 h and 30 min. Table A12 illustrates these results.

The healthy participants sleep on average 18 min more than those with mild disease.

Appendix C.2. Inferential Analysis (PROs vs. TechROs)

We depict the significant correlations between PROs and TechROs for the questionnaires assessing physical activity (Table A13), social support (Table A14), depression and anxiety (Table A15), Mediterranean nutrition (Table A16), nutrition (Table A17), memory (Table A18), sleep (Table A19), and health-related Quality of Life (Table A20). In all tables of this part, we highlight the significant correlations at $r_S \geq 0.5$.

Table A13. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Physical Activity on the IPAQ scale (columns).

		PRO	Work walking activity			Work Moderate Activity			Work Vigorous Activity			Work Total Activity			Transport Walking Activity			Transport Bicycle Activity			Transport Total Activity			Domestic Moderate Activity			Garden Moderate Activity			Garden Vigorous Activity			Domestic+Garden Total Activity			Leisure Walking Activity			Leisure Moderate Activity			Leisure Vigorous Activity			Leisure Total Activity		
Family	**TechRO**	**Health**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**
Raw	Energy		0.4	0.3	0.6	0.5	0.4		0.5	0.5		0.4	0.4					0.4		0.5				0.6	0.4–	0.5–				0.7–	0.4	0.5		0.3–			0.8						0.7			0.6
	Steps		0.5	0.5	0.7	0.4		0.5	0.3			0.4	0.5	0.6	0.4	0.5		0.4		0.5	0.4	0.4		0.5	0.4	0.5	0.6–		0.4	0.7–	0.3	0.5			0.5	0.6–	0.4	0.3	0.8	0.6	0.5	0.7			0.6	0.5	0.4	0.5
	Heart rate					0.4–		0.7–	0.5–		0.7–	0.3–		0.6–				0.3–	0.4–								0.5			0.4–			0.7–							0.3–	0.5–		0.3–			0.4–		
Processed	Sedentary		0.3–	0.5–		0.3–	0.5–		0.3–	0.6–		0.3–	0.5–		0.6–	0.7–		0.4–			0.6–	0.7–					0.5–	0.6		0.5–			0.3–	0.6–	0.5–	0.5–			0.5–	0.7–	0.6–	0.5–	0.7–	0.6–	0.3–			0.5–	0.5–	0.7–	0.6–
	Sedentary+light					0.3–		0.7–	0.3–		0.5–	0.5–						0.3–	0.4–									0.6		0.4–	0.4–	0.5–								0.7		0.3–			0.3–		0.6–	0.3–		0.8–
	Light			0.4					0.3			0.4		0.4		0.6					0.4	0.7	0.7		0.5	0.7	0.6			0.5	0.5–		0.4	0.7		0.5	0.7		0.6	0.7	0.5				0.5–	0.4	0.7	
	Light+fair								0.5	0.5–						0.5					0.5	0.6		0.8	0.6		0.5	0.6				0.5–	0.5	0.6			0.7			0.6			0.6				0.6	
	Fair			0.7–		0.3	0.8–	0.8	0.3	0.6–	0.6	0.3	0.8–		0.3			0.4–	0.6–	0.7				0.6–			0.6–						0.7												0.3		
	Fair+vigorous		0.3		0.6	0.4		0.7	0.5		0.7	0.3		0.6			0.6					0.6	0.6–				0.4–	0.6–					0.7	0.5–								0.3		0.8			
	Vigorous		0.3		0.3	0.7	0.4		0.7	0.3		0.3		0.6	0.6–						0.6			0.4–	0.4–		0.5–	0.6–					0.7	0.6–	0.4–									0.6			
	Active		0.3		0.7	0.3		0.6	0.8			0.3		0.6	0.3	0.5		0.4		0.7	0.4	0.6		0.6				0.6					0.5			0.5		0.6	0.5	0.6	0.6	0.3			0.6	0.6	0.6
	Sleep							0.3										0.5	0.4		0.6	0.6								0.4				0.6–	0.3		0.6					0.5–	0.5			0.6
CLR PA	Sedentary		0.5–		0.7–	0.5–		0.8–	0.6–	0.7–	0.7–	0.5–		0.7–	0.4–			0.5–		0.6–	0.3–	0.7–					0.3	0.5		0.4–		0.7–						0.5–	0.5–		0.3–		0.8–	0.4–		0.5–
	Light			0.5	0.8–	0.5			0.5	0.8–		0.5		0.5	0.6			0.4			0.5	0.5		0.5	0.5		0.5	0.6					0.8–	0.6	0.6				0.4		0.3						
	Fair		0.4		0.6				0.6–	0.3					0.3			0.6	0.4					0.5			0.7	0.6		0.6–			0.5–	0.7											0.3–		0.6–
	Vigorous		0.3	0.5		0.6			0.6			0.6			0.4	0.4					0.6–			0.5	0.6		0.4–	0.6–	0.7–	0.3			0.4–	0.7			0.6	0.5–		0.3		0.8				
CLR PA+S	Sedentary		0.5–			0.5–		0.6–	0.5–		0.7–	0.5–			0.4–	0.5–					0.3–			0.6			0.6			0.5–			0.8–	0.6						0.3–		0.7–	0.4–		0.5–
	Light							0.7–			0.5	0.7–	0.4	0.6				0.6	0.5											0.4	0.5	0.7–											0.8–			
	Fair		0.5	0.4	0.7	0.4	0.4		0.3			0.5	0.4	0.5	0.6–	0.6–		0.6–	0.7–					0.4–	0.5–		0.3–			0.7–			0.5–	0.5–		0.6–			0.3–			0.6–			
	Vigorous		0.5	0.7		0.4	0.6		0.4	0.5		0.4	0.6	0.5	0.7–						0.7–			0.5–	0.6–	0.6–							0.5–	0.4–	0.6–	0.4–	0.4–		0.5				0.6		
	Sleep			0.5		0.6						0.6	0.7–					0.6–	0.7–							0.6–						0.4	0.6													
Raw	Total		1	1	2	1		2	2	1	1		1	2		1					2			4		2	2			2		2	1		1	2			2	1	2	1			2	1		2
Processed	Total			2	2		2	5	2	3	5	1	2	2	1	5		1	1	4	2	4		4	4	2	4	6	1		2	6	3	4	1	3	4	3	3	3	4		1	6	2	4	4
CLR PA	Total		1	2	3	1	2	1	1	3	3	1	2	1	1	1		2		1	2	2		2	3		2	4	1	1		3	2	1	1				2	2			2			2	
CLR PA+S	Total		3	2	1	1	2	2	1	2	2	2	3	3	2	2		3	3		1			2	2	2	1			2	3	2	3	1	1	1			1			1	2	2		1
All Families	Total		5	7	8	3	6	10	6	9	11	4	8	8	4	9		6	4	7	5	6	2	8	11	6	7	10	4	3	7	12	8	7	5	4	4	5	7	7	5	1	1	12	5	4	9

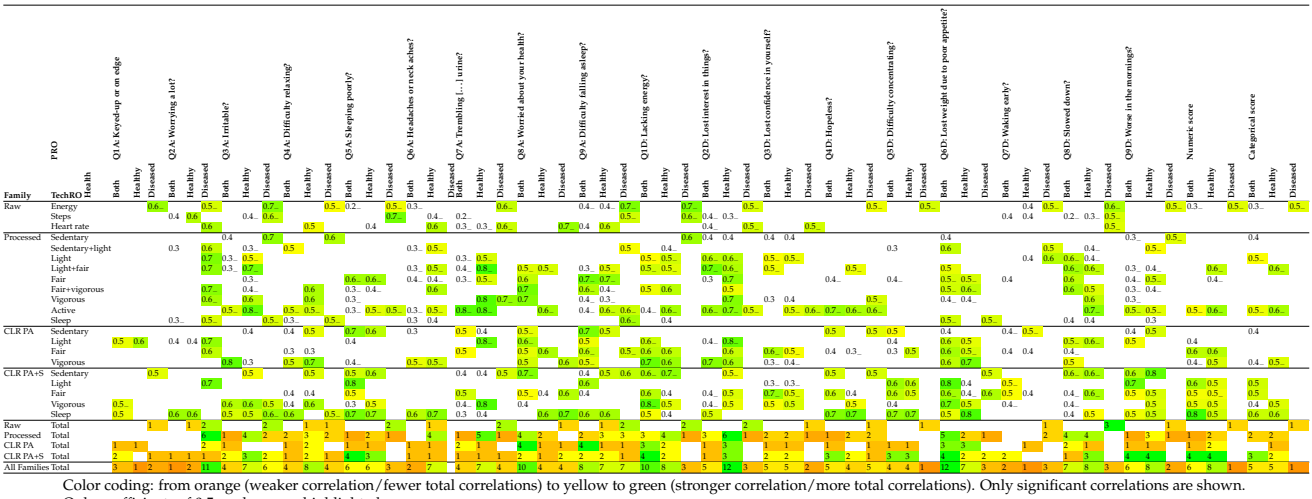
Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A14. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Social Support on the MSPSS scale (columns).

			Q1: Special person: around when in need			Q2: Special person: share joys and sorrows			Q3: Family: tries to help			Q4: Family: gives emotional help and support			Q5: Special person: real source of comfort			Q6: Friends: try to help			Q7: Friends: counted on when things go wrong			Q8: Family: talk about problems			Q9: Friends: share my joys and sorrows			Q10: Special person: cares about feelings			Q11: Family: willing to help make decisions			Q12: Friends: talk about problems			Significant other numeric sub-score			Family numeric sub-score			Friends numeric sub-score			Categorical score			Numeric score		
Family	**TechRO**	**Health**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**
Raw	Energy																																																	
	Steps		0.4	0.4			0.4		0.6	0.6			0.3		0.4	0.5		0.5	0.6	0.5_	0.5	0.6	0.8_		0.4		0.4	0.6	0.7_		0.4		0.5	0.5		0.4	0.7	0.8_	0.3	0.4		0.4	0.4		0.5	0.6	0.6_	0.4	0.5		0.5	0.5	
	Heart rate							0.5_											0.4_			0.4_																										0.3					
Processed	Sedentary		0.5_	0.4_	0.6_	0.5_	0.5_	0.5_	0.5_	0.5_	0.7_	0.5_	0.4_	0.8_	0.6_	0.5_	0.7_	0.6_	0.6_		0.4_	0.5_		0.4_	0.4_		0.4_	0.6_	0.6	0.4_	0.5_		0.7_	0.7_		0.3_	0.6_	0.6	0.5_	0.5_	0.7_	0.6_	0.5_		0.4_	0.5_		0.6_	0.5_		0.6_	0.5_	
	Sedentary+light				0.5_	0.3_	0.3_	0.5_	0.3_						0.5_	0.3_													0.3_	0.3_	0.6_	0.5_								0.3_		0.6						0.4_					
	Light		0.4	0.5		0.4	0.5	0.5_	0.5	0.6		0.4	0.6		0.5	0.6		0.6	0.7	0.7_	0.6	0.7_	0.6_	0.6	0.6		0.6	0.8		0.5	0.5		0.6	0.7		0.6	0.8		0.4	0.6		0.6	0.6		0.7	0.7	0.5_	0.6	0.7		0.6	0.7	0.6_
	Light+fair		0.3	0.5	0.6_	0.4	0.5	0.7_	0.5	0.5			0.6			0.5		0.6	0.7	0.7_	0.6	0.7	0.6_	0.6	0.6		0.7	0.7			0.6	0.5_	0.5	0.6		0.7	0.7		0.4	0.6	0.5_	0.6	0.7		0.7	0.7		0.7	0.7		0.6	0.8	0.6_
	Fair			0.9_			0.9_		0.3_	0.8_			0.7_			0.9_			0.7_			0.7_		0.6_	0.8_			0.4		0.4_	0.8_			0.6_			0.7_			0.9_			0.6_			0.4		0.4_	0.7_			0.7_	
	Fair+vigorous			0.6_			0.7_			0.6_						0.6_			0.7_			0.6_			0.4_	0.5_			0.6_			0.7_			0.6_			0.6_			0.6_			0.7_			0.5_			0.6_		
	Vigorous				0.3_	0.4_										0.6_									0.4_	0.4_			0.7_			0.7_			0.3_	0.7_			0.5_			0.5_			0.6_			0.5_	0.6		0.6_	0.5
	Active		0.6	0.6		0.5	0.6		0.4	0.6		0.5	0.6		0.5	0.6		0.6	0.7	0.5_	0.6	0.5	0.7_	0.5	0.6		0.4	0.6	0.6_	0.5	0.6		0.6	0.7		0.3	0.6	0.7_	0.6	0.6		0.6	0.7		0.6	0.7	0.5_	0.7	0.8		0.6	0.7	
	Sleep				0.5			0.5			0.6			0.6	0.4		0.6													0.5	0.4			0.6					0.3_	0.7											0.6
CLR PA	Sedentary							0.3_																								0.3_																0.3_		0.5_	0.3_		
	Light		0.3	0.4		0.3	0.6		0.4	0.6		0.3	0.4		0.4	0.4			0.5			0.5	0.6_	0.5	0.4		0.5	0.6	0.6_	0.4	0.5		0.3	0.4		0.5	0.5		0.3	0.4		0.3	0.4		0.5	0.5		0.6	0.8		0.3	0.5	
	Fair		0.3	0.5_		0.3_	0.4_		0.4	0.5			0.5_					0.4			0.5			0.7_			0.6	0.5		0.4_	0.4_		0.6	0.8		0.6	0.6		0.3_	0.4_		0.7	0.6		0.5					0.6	0.5	
	Vigorous		0.6_			0.7	0.4		0.7	0.6			0.4		0.7	0.4		0.4	0.5		0.5	0.4		0.6	0.6	0.6	0.4			0.7	0.4		0.6		0.6	0.3	0.4		0.6_		0.6	0.3		0.6	0.6_	0.5	0.7	0.4		0.5	0.4	
CLR PA+S	Sedentary			0.7	0.7_					0.5	0.5_		0.4												0.6		0.4	0.5			0.5	0.6_	0.4_	0.5			0.4_			0.7	0.6_		0.5						0.6_		0.6	0.6_
	Light			0.4		0.3	0.5		0.3	0.5		0.3	0.5		0.4	0.4		0.4	0.5		0.5	0.5		0.6	0.5		0.4	0.6		0.5	0.5		0.4	0.4		0.4	0.5		0.4	0.4		0.4	0.5		0.3	0.4		0.4	0.5		0.4	0.5	
	Fair		0.5	0.5		0.7	0.7		0.8	0.9		0.7	0.4		0.7	0.8		0.3	0.6		0.5_	0.6		0.8	0.8		0.5_			0.7	0.7		0.6	0.6		0.4_	0.5		0.8	0.8		0.6	0.6		0.5	0.5	0.6_	0.7	0.8	
	Vigorous			0.6		0.8	0.6		0.7	0.6			0.4		0.8	0.6		0.3	0.6_		0.3	0.6_		0.8	0.6		0.3	0.5_	0.6_	0.8	0.6		0.5	0.5_		0.4_	0.5		0.4			0.3	0.5_	0.5	0.4	0.6_		0.4	0.4	
	Sleep																0.6	0.3_	0.5_		0.5_	0.4_					0.3_	0.5_					0.6	0.4_		0.3_	0.4_								0.5_			0.3_	0.4_				
Raw	Total							1	1	1						1		1	1	1	1	1	1				1	1					1								1			1	1	1		1		1	1	
Processed	Total		2	5	4	2	6	5	3	6	2	2	4	2	4	6	2	4	7	3	3	6	3	4	5		2	6	2	2	7	3	5	6	1	2	7	2	2	7	4	4	7		3	6	2	4	7	1	4	7	4
CLR PA	Total		1	1		1	1		1	3			1		1			2			2	1	1	3	1	1	2	1	1	1		2	1	1	2	2					1	1	1	3	2		2	1	1	2	2	
CLR PA+S	Total		1	3	1	2	3		2	4	1	1	1		2	2	1	4			3	3		3	4	1	1	4	1	3	4	1	2	2	2		2	3			3	1	1	3		1	3	1	1	2	2	1	3	1
All Families	Total		4	9	5	5	10	6	7	14	3	3	6	2	7	9	3	5	14	4	9	11	5	10	10	1	5	13	5	6	12	5	10	10	2	6	13	3	3	10	6	6	11	1	8	12	4	7	11	4	8	13	5

Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A15. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Depression and Anxiety on the GADS scale (columns).



Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A16. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Mediterranean Nutrition on the PREDIMED scale (columns).

		PRO	Q2: Olive oil use			Q3: Vegetables use			Q4: Fruits use			Q5: Red meat, hamburger, or meat use			Q6: Butter, margarine, or cream use			Q7: Sweet/carbonated beverage use			Q8: Wine use			Q9: Legumes use			Q10: Fish or shellfish use			Q11: Commercial sweets or pastries			Q12: Nuts use			Q14: Sofrito use			Numeric score			Categorical score		
		Health	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased
Family	**TechRO**																																											
Raw	Energy								0.4	0.4					0.8						0.4	0.4																						0.6
	Steps	0.3			0.4	0.4		0.4	0.5						0.3	0.4					0.5	0.4											0.5	0.6					0.4	0.4	0.5	0.4		
	Heart rate	0.4	0.5					0.4							0.6	0.4					0.4			0.4		0.6	0.3			0.3			0.8	0.6	0.4				0.5	0.4	0.5	0.4		
Processed	Sedentary				0.5			0.5	0.4									0.5	0.6		0.3	0.6		0.4																0.6			0.6		
	Sedentary+light	0.3	0.5		0.4			0.5	0.6			0.5		0.4											0.4		0.4												0.3					
	Light					0.5												0.4			0.5	0.6					0.3						0.5	0.4				0.6	0.6		0.5	0.5		
	Light+fair				0.4													0.4	0.5		0.4	0.6					0.3						0.3											
	Fair	0.4							0.7		0.4	0.6									0.4												0.9			0.5	0.6			0.4	0.6			
	Fair+vigorous								0.6																											0.6			0.7			0.7		
	Vigorous														0.6																								0.8			0.7		
	Active				0.4	0.5												0.5	0.5		0.6			0.6						0.4														
	Sleep				0.4			0.4	0.4		0.5	0.5			0.4																								0.3			0.3		
CLR PA	Sedentary								0.6																								0.5	0.7		0.5			0.6			0.7	0.6	
	Light		0.7								0.6	0.5																	0.6			0.4						0.8			0.7		
	Fair				0.6	0.8		0.4	0.6		0.4				0.6			0.4					0.5						0.5										0.6					
	Vigorous							0.5							0.4	0.5		0.4	0.5		0.4						0.6	0.7		0.5									0.5	0.7		0.5	0.6	
CLR PA+S	Sedentary																										0.6			0.6			0.6	0.6	0.6	0.6		0.6			0.6	0.6	
	Light			0.6				0.6	0.5		0.6	0.5									0.6			0.4			0.6			0.6			0.8			0.6				0.6			
	Fair				0.4	0.8		0.5			0.6				0.6									0.5	0.5		0.6									0.7	0.7							
	Vigorous	0.6	0.6		0.5	0.4					0.4					0.6		0.4											0.7							0.7	0.6		0.4			0.5		
	Sleep																				0.7	0.7		0.6			0.6	0.6								0.6							
Raw	Total			1					1					1	1						1					1						1	2	1				1			2		1
Processed	Total			1	1	2		2	3		1	3			1			2	3		2	3		1									2			2	1		3	2		4	2	
CLR PA	Total	1			1	1		1	2		1	1			1	1			1						1		1	1		3			1	1			1		4	1		3	2	
CLR PA+S	Total	1	2		1	1		2	1		2	1			1	1					2	1		2	1		4	1		3			2	1		5	3		1	1		1	2	
All Families	Total	2	4		3	4		5	7		4	5		1	4	2		2	4		5	4		3	2	1	5	2		6		1	7	3		7	5		9	4		10	6	1

Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A17. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Nutrition on the SelfMNA scale (columns).

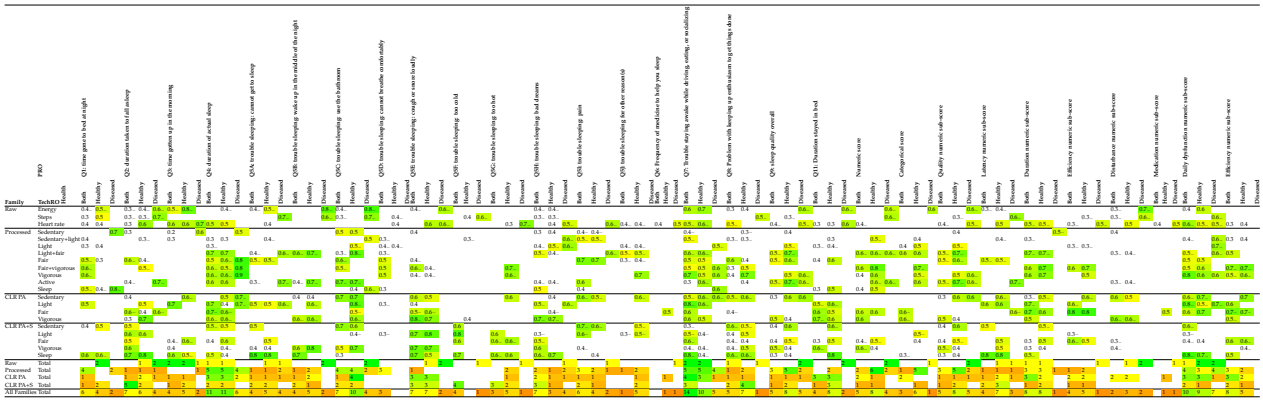
	PRO	Q1: Food Intake Declined			Q2: Weight Lost			Q4: Stressed or Severely Ill			Q5: Dementia and/or Severe Sadness			Numeric Score			Categorical Score		
Family	**TechRO** **Health**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**	**Both**	**Healthy**	**Diseased**
Raw	Energy		0.5_			0.5_			0.4										
	Steps																		
	Heart rate																		
Processed	Sedentary		0.5			0.6			0.8_					0.4					
	Sedentary+light		0.5			0.6													
	Light																	0.5_	
	Light+fair																0.5_		
	Fair																0.6		
	Fair+vigorous																0.4		
	Vigorous													0.4			0.4		
	Active					0.5_			0.6										
	Sleep								0.5					0.4			0.4		
CLR PA	Sedentary																0.6_	0.6_	
	Light																0.6_		
	Fair		0.5_			0.5_									0.6				
	Vigorous														0.7		0.4	0.5	
CLR PA+S	Sedentary																0.6_		
	Light																	0.5_	
	Fair														0.7_				
	Vigorous		0.7_			0.7_													
	Sleep		0.8_			0.8_			0.6								0.6	0.5	
Raw	Total		1			1													
Processed	Total		2			3			3								2	1	
CLR PA	Total		1			1									2		2	2	
CLR PA+S	Total		2			2			1						1		2	2	
All Families	Total		6			7			4						3		6	5	

Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A18. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Memory on the MFE scale (columns).

Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A19. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Sleep on the PSQI scale (columns).



Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

Table A20. Correlation coefficient (Spearman r_S) between TechROs from Fitbit (rows) and PROs of Health-Related Quality of Life on the EQ-5D-3L scale (columns).

		PRO	Q1: Mobility			Q2: Self-Care			Q3: Usual Activities			Q4: Pain/Discomfort			Q5: Anxiety/Depression			Health Score			Mobility Numeric Score			Self-Care Numeric Score			Usual aCtivities Numeric Score			Pain/Discomfort Numeric Score			Anxiety/Depression Numeric Score		
		Health	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased	Both	Healthy	Diseased
Family	**TechRO**																																		
Raw	Energy													0.5–			0.6–												0.5–					0.6–
	Steps		0.4–	0.4–	0.5–												0.5–				0.4–	0.4–	0.5–											0.5–
	Heart rate											0.3–						0.3–	0.5–											0.3–					
Processed	Sedentary		0.4	0.5								0.5	0.6		0.5	0.4	0.8	0.3–	0.3–		0.4	0.5							0.5	0.6		0.5	0.4	0.8
	Sedentary+light		0.3										0.5		0.5	0.4					0.3									0.5		0.5	0.4	
	Light		0.3–	0.3–										0.6	0.4–	0.4–			0.4		0.3–	0.3–								0.6	0.4–	0.4–		
	Light+fair																																		
	Fair											0.3–	0.5–	0.6–		0.5		0.5–	0.4											0.3–	0.5–	0.6–		0.5	
	Fair+vigorous												0.5–					0.5–													0.5–				
	Vigorous															0.4		0.6–	0.5–														0.4		
	Active		0.3–													0.5	0.6–	0.7–			0.3–												0.5	0.6–	
	Sleep											0.5–			0.5–	0.4–	0.5–	0.4												0.5–			0.5–	0.4–	0.5–
CLR PA	Sedentary		0.3										0.6		0.4	0.5–		0.6			0.3									0.6		0.4	0.5–	
	Light														0.3–			0.4	0.4														0.3–		
	Fair											0.5–	0.6–						0.5–	0.6–										0.5–	0.6–				
	Vigorous																	0.3–		0.6–															
CLR PA+S	Sedentary												0.6					0.6												0.6				
	Light															0.4–			0.4														0.4–		
	Fair											0.6–						0.6	0.6–											0.6–					
	Vigorous											0.5						0.3–	0.6–											0.5					
	Sleep											0.5–			0.3–			0.5												0.5–			0.3–		
Raw	Total				1									1			2		1				1								1				2
Processed	Total			1								2	4	2	3	2	3	4	1			1							2	4	2	3	2	3
CLR PA	Total											1	2			1		1	1	2									1	2			1	
CLR PA+S	Total											3	1					3	2										3	1				
All Families	Total			1	1							6	7	3	3	3	5	8	5	2		1	1						6	7	3	3	3	5

Color coding: from orange (weaker correlation/fewer total correlations) to yellow to green (stronger correlation/more total correlations). Only significant correlations are shown. Only coefficients of 0.5 or above are highlighted.

References

1. Naghavi, M.; Abajobir, A.A.; Abbafati, C.; Abbas, K.M.; Abd-Allah, F.; Abera, S.F.; Aboyans, V.; Adetokunboh, O.; Afshin, A.; Agrawal, A.; et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **2017**, *390*, 1151–1210.
2. WHO Regional Office for Europe. *Noncommunicable Diseases*; WHO Regional Office for Europe: Copenhagen, Denmark, 2020.
3. Anderson, G. Responding to the Growing Cost and Prevalence of People with Multiple Chronic Conditions. Ph.D. Thesis, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, 2011.
4. United States Office of the Assistant Secretary for Health. *Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention*; United States Public Health Service: Washington DC, USA, 1979.
5. United Nations Department of Economic and Social Affairs *World Population Prospects 2019: Highlights*; United Nations Department for Economic and Social Affairs: New York, NY, USA, 2019.
6. Mayo, N.E.; Figueiredo, S.; Ahmed, S.; Bartlett, S.J. Montréal Accord on Patient-Reported Outcomes (PROs) use series–Paper 2: Terminology proposed to measure what matters in health. *J. Clin. Epidemiol.* **2017**, *89*, 119–124.
7. Sims, J.; Smith, F.; Duffy, A.; Hilton, S. The vagaries of self-reports of physical activity: A problem revisited and addressed in a study of exercise promotion in the over 65 s in general practice. *Fam. Pract.* **1999**, *16*, 152–157.
8. Anderson, K.N.; Catt, M.; Collerton, J.; Davies, K.; von Zglinicki, T.; Kirkwood, T.B.; Jagger, C. Assessment of sleep and circadian rhythm disorders in the very old: the Newcastle 85+ Cohort Study. *Age Ageing* **2014**, *43*, 57–63.

9. Van Den Berg, J.F.; Van Rooij, F.J.; Vos, H.; Tulen, J.H.; Hofman, A.; Miedema, H.M.; Neven, A.K.; Tiemeier, H. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. *J. Sleep Res.* **2008**, *17*, 295–302.
10. Wac, K. From quantified self to quality of life. In *Digital Health*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 83–108.
11. Sylvia, L.G.; Bernstein, E.E.; Hubbard, J.L.; Keating, L.; Anderson, E.J. Practical guide to measuring physical activity. *J. Acad. Nutr. Diet.* **2014**, *114*, 199–208.
12. Ferguson, T.; Rowlands, A.V.; Olds, T.; Maher, C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: A cross-sectional study. *Int. J. Behav. Nutr. Phys. Act.* **2015**, *12*, 42.
13. International Data Corporation. *IDC Forecasts Sustained Double-Digit Growth for Wearable Devices Led by Steady Adoption of Smartwatches*; International Data Corporation: San Mateo, CA, USA, 2019.
14. Dominick, G.M.; Winfree, K.N.; Pohlig, R.T.; Papas, M.A. Physical activity assessment between consumer-and research-grade accelerometers: A comparative study in free-living conditions. *JMIR mHealth uHealth* **2016**, *4*, e110.
15. McMahon, S.K.; Lewis, B.; Oakes, M.; Guan, W.; Wyman, J.F.; Rothman, A.J. Older adults' experiences using a commercially available monitor to self-track their physical activity. *JMIR mHealth uHealth* **2016**, *4*, e35.
16. Martinez, G.J.; Mattingly, S.M.; Mirjafari, S.; Nepal, S.K.; Campbell, A.T.; Dey, A.K.; Striegel, A.D. On the Quality of Real-world Wearable Data in a Longitudinal Study of Information Workers. In Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Austin, TX, USA, 23–27 March 2020; pp. 1–6.
17. Adam Noah, J.; Spierer, D.K.; Gu, J.; Bronner, S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. *J. Med Eng. Technol.* **2013**, *37*, 456–462.
18. Imboden, M.T.; Nelson, M.B.; Kaminsky, L.A.; Montoye, A.H. Comparison of four Fitbit and Jawbone activity monitors with a research-grade ActiGraph accelerometer for estimating physical activity and energy expenditure. *Br. J. Sport. Med.* **2018**, *52*, 844–850.
19. Van Camp, C. Accuracy of Predicted Energy Expenditure from a Fitbit Inspire HR Activity Monitor during Short-and Long-Duration Exercise. Master's Thesis, Eastern Michigan University, Ypsilanti, MI, USA, 2020.
20. CoME Consortium. Caregiver & ME (CoME): Application for Healthier Self-Management of Seniors. 2017. Available online: http://come-aal.eu/ (accessed on 1 August 2020).
21. De Vet, H.; Terwee, C.; Mokkink, L.; Knol, D. Concepts, theories and models, and types of measurements. In *Measurement in Medicine: A Practical Guide*; Cambridge University Press: Cambridge, UK, 2011; pp. 7–29.
22. Livingston, G.; Sommerlad, A.; Orgeta, V.; Costafreda, S.G.; Huntley, J.; Ames, D.; Ballard, C.; Banerjee, S.; Burns, A.; Cohen-Mansfield, J.; et al. Dementia prevention, intervention, and care. *Lancet* **2017**, *390*, 2673–2734.
23. Sindi, S.; Kåreholt, I.; Johansson, L.; Skoog, J.; Sjöberg, L.; Wang, H.X.; Johansson, B.; Fratiglioni, L.; Soininen, H.; Solomon, A.; et al. Sleep disturbances and dementia risk: A multicenter study. *Alzheimer's Dement.* **2018**, *14*, 1235–1242.
24. Chen, J.C.; Espeland, M.A.; Brunner, R.L.; Lovato, L.C.; Wallace, R.B.; Leng, X.; Phillips, L.S.; Robinson, J.G.; Kotchen, J.M.; Johnson, K.C.; others. Sleep duration, cognitive decline, and dementia risk in older women. *Alzheimer's Dement.* **2016**, *12*, 21–33.
25. Benedict, C.; Byberg, L.; Cedernaes, J.; Hogenkamp, P.S.; Giedratis, V.; Kilander, L.; Lind, L.; Lannfelt, L.; Schiöth, H.B. Self-reported sleep disturbance is associated with Alzheimer's disease risk in men. *Alzheimer's Dementia* **2015**, *11*, 1090–1097.
26. The IPAQ Group. International Physical Activity Questionnaire. 2002. Available online: https://sites.google.com/site/theipaq/ (accessed on 1 August 2020).
27. Zimet, G.D.; Dahlem, N.W.; Zimet, S.G.; Farley, G.K. The multidimensional scale of perceived social support. *J. Personal. Assess.* **1988**, *52*, 30–41.
28. Goldberg, D.; Bridges, K.; Duncan-Jones, P.; Grayson, D. Detecting anxiety and depression in general medical settings. *Br. Med J.* **1988**, *297*, 897–899.
29. Estruch, R.; Ros, E.; Salas-Salvadó, J.; Covas, M.I.; Corella, D.; Arós, F.; Gómez-Gracia, E.; Ruiz-Gutiérrez, V.; Fiol, M.; Lapetra, J.; et al. Primary prevention of cardiovascular disease with a Mediterranean diet. *N. Engl. J. Med.* **2013**, *368*, 1279–1290.

30. Estruch, R.; Ros, E.; Salas-Salvadó, J.; Covas, M.I.; Corella, D.; Arós, F.; Gómez-Gracia, E.; Ruiz-Gutiérrez, V.; Fiol, M.; Lapetra, J.; et al. Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts. *N. Engl. J. Med.* **2018**, *378*, e34.
31. Nestlé Nutrition Institute. Mini Nutritional Assessment (Self MNA®) for Adults 65 Years of Age and Older. 2012. Available online: https://www.mna-elderly.com/ (accessed on 1 August 2020).
32. Sunderland, A.; Watts, K.; Baddeley, A.D.; Harris, J.E. Subjective memory assessment and test performance in elderly adults. *J. Gerontol.* **1986**, *41*, 376–384.
33. Carpenter, J.S.; Andrykowski, M.A. Psychometric evaluation of the Pittsburgh sleep quality index. *J. Psychosom. Res.* **1998**, *45*, 5–13.
34. EuroQol Research Foundation. EQ-5D-3L. 2020. Available online: https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/ (accessed on 1 August 2020).
35. Fitbit, Inc. Fitbit Official Website. 2020. Available online: https://www.fitbit.com/dk/home (accessed on 29 October 2020).
36. Garriguet, D.; Tremblay, S.; Colley, R.C. Comparison of physical activity adult questionnaire results with accelerometer data. *Health Rep.* **2015**, *26*, 11.
37. Brewer, W.; Swanson, B.T.; Ortiz, A. Validity of Fitbit's active minutes as compared with a research-grade accelerometer and self-reported measures. *BMJ Open Sport Exerc. Med.* **2017**, *3*, e000254.
38. Rossi, A.; Frechette, L.; Miller, D.; Miller, E.; Friel, C.; Van Arsdale, A.; Lin, J.; Shankar, V.; Kuo, D.Y.; Nevadunsky, N.S. Acceptability and feasibility of a Fitbit physical activity monitor for endometrial cancer survivors. *Gynecol. Oncol.* **2018**, *149*, 470–475.
39. Meyer, J.; von Holdt, K.; Beck, E.; Brandes, M.; Pischke, C.; Voelcker-Rehage, C. Toy or tool? Activity trackers for the assessment of physical activity in the wild. In Proceedings of the 2019 IEEE International Conference on Healthcare Informatics (ICHI), Xi'an, China, 10–13 June 2019; pp. 1–9.
40. Santomas, K. *Passive Monitoring of Physical Activity Using a Fitbit Charge 3 in Individuals at Risk for Depression*; Technical Report; University of Massachussetts Lowell: Lowell, MA, USA, 2020.
41. Troxel, W.M.; Buysse, D.J.; Monk, T.H.; Begley, A.; Hall, M. Does social support differentially affect sleep in older adults with versus without insomnia? *J. Psychosom. Res.* **2010**, *69*, 459–466.
42. Carlson, J.A.; Sallis, J.F.; Conway, T.L.; Saelens, B.E.; Frank, L.D.; Kerr, J.; Cain, K.L.; King, A.C. Interactions between psychosocial and built environment factors in explaining older adults' physical activity. *Prev. Med.* **2012**, *54*, 68–73.
43. McMurdo, M.E.; Argo, I.; Crombie, I.K.; Feng, Z.; Sniehotta, F.F.; Vadiveloo, T.; Witham, M.D.; Donnan, P.T. Social, environmental and psychological factors associated with objective physical activity levels in the over 65 s. *PLoS ONE* **2012**, *7*, e31878.
44. Ho, E.C.; Hawkley, L.; Dale, W.; Waite, L.; Huisingh-Scheetz, M. Social capital predicts accelerometry-measured physical activity among older adults in the US: A cross-sectional study in the National Social Life, Health, and Aging Project. *BMC Public Health* **2018**, *18*, 804.
45. Arigo, D.; Pasko, K.; Mogle, J.A. Daily relations between social perceptions and physical activity among college women. *Psychol. Sport Exerc.* **2019**, *47*, 101528.
46. Doryab, A.; Villalba, D.K.; Chikersal, P.; Dutcher, J.M.; Tumminia, M.; Liu, X.; Cohen, S.; Creswell, K.; Mankoff, J.; Creswell, J.D.; et al. Identifying Behavioral Phenotypes of Loneliness and Social Isolation with Passive Sensing: Statistical Analysis, Data Mining and Machine Learning of Smartphone and Fitbit Data. *JMIR mHealth uHealth* **2019**, *7*, e13209.
47. Adams, D.J.; Remick, R.A.; Davis, J.C.; Vazirian, S.; Khan, K.M. Exercise as medicine—the use of group medical visits to promote physical activity and treat chronic moderate depression: A preliminary 14-week pre–post study. *BMJ Open Sport Exerc. Med.* **2015**, *1*, e000036.
48. Helgadóttir, B.; Forsell, Y.; Ekblom, Ö. Physical activity patterns of people affected by depressive and anxiety disorders as measured by accelerometers: A cross-sectional study. *PLoS ONE* **2015**, *10*, e0115894.
49. Liau, A.K.; Neihart, M.; Teo, C.T.; Goh, L.S.; Chew, P. A Quasi-experimental study of a fitbit-based self-regulation intervention to improve physical activity, well-being, and mental health. *Cyberpsychol. Behav. Soc. Netw.* **2018**, *21*, 727–734.
50. Vidal Bustamante, C.M.; Rodman, A.M.; Dennison, M.J.; Flournoy, J.C.; Mair, P.; McLaughlin, K.A. Within-person fluctuations in stressful life events, sleep, and anxiety and depression symptoms during adolescence: A multiwave prospective study. *J. Child Psychol. Psychiatry* **2020**, 61, 1116–1125

51. Falck, R.S.; Landry, G.J.; Best, J.R.; Davis, J.C.; Chiu, B.K.; Liu-Ambrose, T. Cross-sectional relationships of physical activity and sedentary behavior with cognitive function in older adults with probable mild cognitive impairment. *Phys. Ther.* **2017**, *97*, 975–984.
52. Giannakos, M.N.; Sharma, K.; Papavlasopoulou, S.; Pappas, I.O.; Kostakos, V. Fitbit for learning: Towards capturing the learning experience using wearable sensing. *Int. J. Hum. Comput. Stud.* **2020**, *136*, 102384.
53. McLaren, B.; Andrews, S.C.; Glikmann-Johnston, Y.; Mercieca, E.C.; Murray, N.W.; Loy, C.; Bellgrove, M.A.; Stout, J.C. Feasibility and initial validation of 'HD-Mobile', a smartphone application for remote self-administration of performance-based cognitive measures in Huntington's disease. *J. Neurol.* **2020**, 1432–1459, DOI: 10.1007/s00415-020-10169-y.
54. Sano, A.; Phillips, A.J.; Amy, Z.Y.; McHill, A.W.; Taylor, S.; Jaques, N.; Czeisler, C.A.; Klerman, E.B.; Picard, R.W. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. In Proceedings of the 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Cambridge, MA, USA, 9–12 June 2015; pp. 1–6.
55. Peach, H.D.; Gaultney, J.F.; Ruggiero, A.R. Direct and indirect associations of sleep knowledge and attitudes with objective and subjective sleep duration and quality via sleep hygiene. *J. Prim. Prev.* **2018**, *39*, 555–570.
56. Thota, D. Evaluating the Relationship Between Fitbit Sleep Data and Self-Reported Mood, Sleep, and Environmental Contextual Factors in Healthy Adults: Pilot Observational Cohort Study. *JMIR Form. Res.* **2020**, *4*, e18086.
57. Lohne-Seiler, H.; Hansen, B.H.; Kolle, E.; Anderssen, S.A. Accelerometer-determined physical activity and self-reported health in a population of older adults (65–85 years): A cross-sectional study. *BMC Public Health* **2014**, *14*, 284.
58. Mobbs, R.J.; Phan, K.; Maharaj, M.; Rao, P.J. Physical activity measured with accelerometer and self-rated disability in lumbar spine surgery: A prospective study. *Glob. Spine J.* **2016**, *6*, 459–464.
59. Gresham, G.; Hendifar, A.E.; Spiegel, B.; Neeman, E.; Tuli, R.; Rimel, B.; Figlin, R.A.; Meinert, C.L.; Piantadosi, S.; Shinde, A.M. Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients. *NPJ Digit. Med.* **2018**, *1*, 27.
60. Sasaki, S.; Kanai, M.; Shinoda, T.; Morita, H.; Shimada, S.; Izawa, K.P. Relation between health utility score and physical activity in community-dwelling ambulatory patients with stroke: A preliminary cross-sectional study. *Top. Stroke Rehabil.* **2018**, *25*, 475–479.
61. Twiggs, J.; Salmon, L.; Kolos, E.; Bogue, E.; Miles, B.; Roe, J. Measurement of physical activity in the pre-and early post-operative period after total knee arthroplasty for Osteoarthritis using a Fitbit Flex device. *Med. Eng. Phys.* **2018**, *51*, 31–40.
62. Laghouila, S.; Manea, V.; Estrada, V.; Wac, K. Digital Health Tools for Chronic Illness and Dementia Risk Assessment in Older Adults. In *Annals of Behavioral Medicine*; Oxford Univ Press Inc.: Oxford, UK, 2018; Volume 52, pp. S291–S291.
63. Fitbit, Inc. *Fitbit Developer API—Daily Activity Summary*; Fitbit, Inc.: San Francisco, CA, USA, 2019.
64. Rosenberger, M.E.; Fulton, J.E.; Buman, M.P.; Troiano, R.P.; Grandner, M.A.; Buchner, D.M.; Haskell, W.L. The 24-H Activity Cycle: A New Paradigm for Physical Activity. *Med. Sci. Sport. Exerc.* **2019**, *51*, 454–464.
65. Aitchison, J. The statistical analysis of compositional data. *J. R. Stat. Soc. Ser. B Methodol.* **1982**, *44*, 139–160.
66. Fitbit, Inc. *Fitbit User Forums: Lightly Active vs. Active vs. Very Active*; Fitbit, Inc.: San Francisco, CA, USA, 2019.
67. Fitbit, Inc. *Fitbit User Forums: How Activity Categories Are Determined*; Fitbit, Inc.: San Francisco, CA, USA, 2019.
68. Fitbit, Inc. *What are Active Minutes?*; Fitbit, Inc.: San Francisco, CA, USA, 2019.
69. Smith, G.L.; Banting, L.; Eime, R.; O'Sullivan, G.; Van Uffelen, J.G. The association between social support and physical activity in older adults: A systematic review. *Int. J. Behav. Nutr. Phys. Act.* **2017**, *14*, 56.
70. Spearman, C. *The Proof and Measurement of Association between Two Things*; Appleton-Century-Crofts: New York, NY, USA, 1961.
71. The SciPy community. SciPy 1.5.0 Release Notes. 2020. Available online: https://docs.scipy.org/doc/scipy/reference/release.1.5.0.html (accessed on 1 August 2020).

72. Wanner, M.; Probst-Hensch, N.; Kriemler, S.; Meier, F.; Autenrieth, C.; Martin, B.W. Validation of the long international physical activity questionnaire: influence of age and language region. *Prev. Med. Rep.* **2016**, *3*, 250–256.
73. Kent de Grey, R.G.; Uchino, B.N.; Trettevik, R.; Cronan, S.; Hogan, J.N. Social support and sleep: A meta-analysis. *Health Psychol.* **2018**, *37*, 787.
74. Monroe, C.M.; Thompson, D.L.; Bassett, D.R., Jr.; Fitzhugh, E.C.; Raynor, H.A. Usability of mobile phones in physical activity–related research: A systematic review. *Am. J. Health Educ.* **2015**, *46*, 196–206.
75. Feng, Y.; Chang, C.K.; Chang, H. An adl recognition system on smart phone. In *International Conference on Smart Homes and Health Telematics*; Springer: Cham, Switzerland, 2016; pp. 148–158.
76. Feng, Y.; Chang, C.K.; Ming, H. Recognizing activities of daily living to improve well-being. *IT Prof.* **2017**, *19*, 31–37.
77. Simon, E.B.; Walker, M.P. Sleep loss causes social withdrawal and loneliness. *Nat. Commun.* **2018**, *9*, 1–9.
78. Wu, J.; Feng, Y.; Sun, P. Sensor fusion for recognition of activities of daily living. *Sensors* **2018**, *18*, 4029.
79. Hebden, L.; Cook, A.; Van Der Ploeg, H.P.; Allman-Farinelli, M. Development of smartphone applications for nutrition and physical activity behavior change. *JMIR Res. Protoc.* **2012**, *1*, e9.
80. Goldberg, R. Escala de Goldberg. 1993. Available online: http://www.sld.cu/galerias/pdf/sitios/rehabilitacion/goldberg-depresion.pdf (accessed on 1 August 2020).
81. EuroQol Research Foundation. Available EQ-5D Versions. 2020. Available online: https://euroqol.org/eq-5d-instruments/all-eq-5d-versions/ (accessed on 1 August 2020).
82. Carrasco, P.M.; Peña, M.M.; Sueiro, M.J. The memory failures of everyday questionnaire (MFE): Internal consistency and reliability. *Span. J. Psychol.* **2012**, *15*, 768.
83. Nestlé Nutrition Institute. Mini Nutritional Assessment (Self MNA®) Para Adultos a Partir de 65 Años. 2012. Available online: https://www.mna-elderly.com/forms/Self-MNA_Spanish.pdf (accessed on 1 August 2020).
84. Centro de Investigacion Biomedica en Red. Estudio Predimed. 2020. Available online: http://www.predimed.es/uploads/8/0/5/1/8051451/p14_medas.pdf (accessed on 1 August 2020).
85. The IPAQ Group. Cuestionario Internacional de Actividad Fisica. 2002. Available online: https://sites.google.com/site/theipaq/questionnaire_links/IPAQ_Spanish%28USA%29_self-admin_long.pdf (accessed on 1 August 2020).
86. Universitat de Barcelona. Indice de calidad de sueño de Pittsburgh. 1998. Available online: http://www.ub.edu/psicobiologia/Pmemlleng/images/Index%20de%20Pittsburgh.pdf (accessed on 1 August 2020).
87. American Academy of Pediatrics. Multidimensional Scale of Perceived Social Support (Spanish). 2010. Available online: http://www.heardalliance.org/wp-content/uploads/2011/04/Multidimensional-Scale-of-Perceived-Social-Support-Spanish.pdf (accessed on 1 August 2020).
88. Orsolya, P.Z.; Zoltán, K.M.; Szilvia, J. A Multidimenzionális észlelt társas támogatás kérdőív magyar nyelvű validálása. *Mentálhigiéné Pszichoszomatika* **2017**, *18*, 230–262.
89. The IPAQ Group. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). 2005. Available online: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWRlZTM (accessed on 1 August 2020).
90. Rasch, G. *Probabilistic Models for Some Intelligence and Attainment Tests*; ERIC: Washington, DC, USA, 1993.
91. Inkin. *Wearables: The Comprehensive List of Wearables on the Market*; 2020. Available online: https://www.inkin.com/wearables/ (accessed on 29 October 2020).
92. United States Food and Drug Administration. *FDA Selects Participants for New Digital Health Software Precertification Pilot Program*; United States Food and Drug Administration: Silver Spring, MD, USA, 2017.
93. Kooiman, T.J.; Dontje, M.L.; Sprenger, S.R.; Krijnen, W.P.; van der Schans, C.P.; de Groot, M. Reliability and validity of ten consumer activity trackers. *BMC Sports Sci. Med. Rehabil.* **2015**, *7*, 24.
94. Keill, A. Validity of wearable fitness trackers on sleep measure. *Med. Sci. Sports Exerc.* **2016**, *48*, pp. 10.
95. Lee, J.M.; Byun, W.; Keill, A.; Dinkel, D.; Seo, Y. Comparison of Wearable Trackers' Ability to Estimate Sleep. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1265.
96. Tedesco, S.; Sica, M.; Ancillao, A.; Timmons, S.; Barton, J.; O'Flynn, B. Validity Evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in Free-Living Environments in an Older Adult Cohort. *JMIR mHealth uHealth* **2019**, *7*, e13084.

97. Paul, S.S.; Tiedemann, A.; Hassett, L.M.; Ramsay, E.; Kirkham, C.; Chagpar, S.; Sherrington, C. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. *BMJ Open Sport Exerc. Med.* **2015**, *1*, e000013.
98. Brooke, J. SUS-A quick and dirty usability scale. *Usability Eval. Ind.* **1996**, *189*, 4–7.
99. Steinert, A.; Haesner, M.; Steinhagen-Thiessen, E. Activity-tracking devices for older adults: Comparison and preferences. *Univers. Access Inf. Soc.* **2018**, *17*, 411–419.
100. Cooper, C.; Gross, A.; Brinkman, C.; Pope, R.; Allen, K.; Hastings, S.; Bogen, B.E.; Goode, A. The impact of wearable motion sensing technology on physical activity in older adults. *Exp. Gerontol.* **2018**, *112*, 9–19.
101. Rothman, K.J. No adjustments are needed for multiple comparisons. *Epidemiology* **1990**, *1*, 43–46.

Publication 7: Using Consumer-Friendly Wearables to Associate Patient- and Technology-Reported Physical Activity in Healthy Seniors

Vlad Manea, Allan Berrocal, Katarzyna Wac. *Using Consumer-Friendly Wearables to Associate Patient- and Technology-Reported Physical Activity in Healthy Seniors*. Conference on Mobile Systems and Pervasive Computing (MobiSPC 2020). 8p. DOI: https://doi.org/10.1016/j.procs.2020.07.036. [56]

Abstract A leading risk factor for chronic disease is physical inactivity. In efforts to assess physical activity and inform designs for prevention, health professionals currently use inexpensive, but subjective validated scales, or objective, but expensive research-grade wearables. In the meanwhile, individuals increasingly use affordable consumer-friendly wearable devices that can objectively monitor behaviours while daily life unfolds. However, the relationships between their outcomes and the validated scales are yet to be calibrated. We report our results from a study on 31 seniors from Hungary and Spain (mean age 70.6 $\pm$ 3.2). Our study quantified the relations between physical activity outcomes, as patient-reported through 53 answers (1.71 $\pm$ 0.96 / person) on the International Physical Activity Questionnaire (IPAQ) with a 7-day recall period, and 5615 days (mean 181.1 $\pm$ 179.2 days collected /person) technology-reported by Fitbit Charge 2. The wearables monitored daily life behaviours of physical activity and sleep for long durations (7 to 120 days). We found strong Spearman correlations between light and moderate IPAQ physical activity in the domestic activity domain, and light-fair intensity Fitbit physical activity (e.g., r_S = 0.88, p < 0.005). We also found negative moderate-strong correlations between Fitbit sedentary duration and all IPAQ physical activity domains and intensities (e.g., r_S = 0.64, p < 0.005). We obtained increasingly stronger relationships across all IPAQ domains and Fitbit intensities by monitoring physical activity beyond the scale recall period, quantifying physical activity relative to all activities of the day, and including sleep. Our findings inform the design of longitudinal observations and personalized, focused, and potentially effective interventions for physical activity in seniors.

Keywords observational study, healthy senior, physical activity, statistical correlation, questionnaire, IPAQ, consumer-friendly wearable, Fitbit.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 175 (2020) 245–252

Procedia Computer Science

www.elsevier.com/locate/procedia

The 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC)
August 9-12, 2020, Leuven, Belgium

Using consumer-friendly wearables to correlate patient and technology-reported physical activity in healthy seniors

Vlad Manea[a], Allan Barrocal[b], Katarzyna Wac[a,b,*]

[a] *Quality of Life Technologies Lab, Department of Computer Science, University of Copenhagen, Sigurdsgade 41, Copenhagen 2200, Denmark*
[b] *Quality of Life Technologies Lab, Institute of Service Science, University of Geneva, 7 Route de Drize, Geneva 1227, Switzerland*

Abstract

A leading risk factor for chronic disease is physical inactivity. In efforts to assess physical activity and inform designs for prevention, health professionals currently use inexpensive, but subjective validated scales, or objective, but expensive research-grade wearables. In the meanwhile, individuals increasingly use affordable consumer-friendly wearable devices that can objectively monitor behaviours while daily life unfolds. However, the relationships between their outcomes and the validated scales are yet to be calibrated. We report our results from a study on 31 seniors from Hungary and Spain (mean age 70.6 +/- 3.2). Our study quantified the relations between physical activity outcomes, as patient-reported through 53 answers (1.71 +/- 0.96 / person) on the International Physical Activity Questionnaire (IPAQ) with a 7-day recall period, and 5615 days (mean 181.1 +/- 179.2 days collected / person) technology-reported by Fitbit Charge 2. The wearables monitored daily life behaviours of physical activity and sleep for long durations (7 to 120 days). We found strong Spearman correlations between light and moderate IPAQ physical activity in the domestic activity domain, and light-fair intensity Fitbit physical activity (e.g., $r_S = 0.88$, $p < 0.005$). We also found negative moderate-strong correlations between Fitbit sedentary duration and all IPAQ physical activity domains and intensities (e.g., $r_S = 0.64$, $p < 0.005$). We obtained increasingly stronger relationships across all IPAQ domains and Fitbit intensities by monitoring physical activity beyond the scale recall period, quantifying physical activity relative to all activities of the day, and including sleep. Our findings inform the design of longitudinal observations and personalized, focused, and potentially effective interventions for physical activity in seniors.

Keywords: observational study; healthy senior; physical activity; statistical correlation; questionnaire; IPAQ; consumer-friendly wearable; Fitbit

* Corresponding author. Tel.: +41-22-379-024
E-mail address: katarzyna.wac@unige.ch

10.1016/j.procs.2020.07.036

1. Introduction

Chronic diseases represent a significant share of the burden of disease globally [21]. They are responsible for 86% of all deaths [36]. In Europe, chronic diseases affect over 80% of adults over 65 and incur 70% of the increasing healthcare costs [3]. The most common chronic diseases are cardiovascular, pancreatic, pulmonary, and neoplastic.

Unhealthy lifestyle and behaviours, such as physical inactivity, poor nutrition, and tobacco intake, explain up to 50% of the risk of chronic disease [12]. A leading behaviour of risk is physical inactivity. There is *"overwhelming evidence that proves the notion that reductions in daily physical activity are primary causes of chronic diseases"* [5]. However, *"the evidence is currently insufficiently precise to warrant separate guidelines for each specific disease, but it is strong enough to cover all health outcomes"* [35].

The gold standard for the measurement of physical activity is the subjective patient-reported outcome (PRO, [19]) administered as a questionnaire based on a qualitative scale, statistically validated on a population of interest. However, validated scales for physical activity have the inherent shortcomings of PROs: they are inconvenient, infrequent, memory-biased, socially conditioned, and qualitative.

Research-grade wearables measure physical activity objectively. They provide quantitative, technology-reported physical activity outcomes (TechRO, [19]), such as acceleration and heart rate, and have been clinically validated. However, they are uncomfortable and expensive [30]. Several studies used PROs from validated scales and TechROs from research-grade wearables to quantify the relationships between subjective and objective physical activity [11, 34]. However, their participants wore the devices for a *short-term* period, and without owning the devices.

Consumer-friendly wearables measure continuously, accurately, and objectively quantitative TechROs of physical activity during daily life for long, or *longitudinal* periods [9]. Also, more individuals opt for consumer-friendly wearable devices; the market size will likely double by 2022 [16]. However, the few studies using exclusively consumer-friendly wearables to measure longitudinal physical activity, e.g., [6], focused on younger populations.

To our knowledge, there are no studies aimed at quantifying the relationships between physical activity PROs (obtained from validated scales) and TechROs (collected from consumer-friendly wearables) at different intensities in longitudinal, daily, and free-living conditions for seniors. Our study observed N = 31 healthy seniors along 2017-2019. They provided 53 PRO answers and collected 5615 TechRO days of physical activity and sleep. We included sleep in the TechROs to model the interdependence of active, sedentary, and sleep duration during the day [26].

From over 80 scales that provide physical activity PROs [22], we chose the International Physical Activity Questionnaire (IPAQ). The IPAQ is *"developed to measure health-related physical activity in populations"* [14] and has been validated on seniors. In our study, it was feasible to administer the long (and more detailed) variant of the IPAQ.

From over 200 models of consumer-friendly wearables that provide physical activity TechROs [15], we chose Fitbit (Fitbit, Inc.). Fitbit aims at motivating consumers to *"reach health and fitness goals by tracking activity, exercise, sleep, weight, and more"* [10]. It has been selected for Digital Health Software Precertification by the US FDA [32]. Fitbit monitors daily life behaviours accurately and continuously, operationalizes the critical human factors for prolonged wear by senior end-users, and facilitates reliable behavioural data collection. We selected Fitbit Charge 2, a watch model with a user-friendly display, previously validated with seniors [31].

Our paper is structured as follows. Section 2 reviews related work. Section 3 describes our method. Section 4 foregrounds our results. Section 5 discusses our findings. The last section concludes the paper.

2. Related Work

Validated scales of physical activity and sleep have only moderate validity and reliability [24]. Sims et al. have shown that seniors reporting physical activity overestimate the amount undertaken (N = 20, mean age 72.2, [28]). Also, two studies by Anderson (N = 421, ages 87-89, [4]) and Van Der Berg (N = 69, ages 57-97, [33]) have shown that subjective sleep is less reliable than objective sleep.

Several studies have compared PROs to TechROs by using a validated scale in tandem with a research-grade wearable. For example, Garriguet et al. have found a Spearman $r_S = 0.23$ correlation between PRO and TechRO moderate and vigorous activity. The latter was reported by Actical accelerometers worn for seven days (N = 112, ages 18-79, [11]). Wanner et al. have obtained a Spearman $r_S = 0.41$ correlation between vigorous physical activity from

PROs and TechROs. They used ActiGraph GTX3+ accelerometers worn for eight days (N = 346, mean age 54.6, [34]). These studies have focused on younger populations for a *short-term* period (typically one week).

Other studies have compared TechROs from both research-grade and consumer-friendly wearables by requiring the participants to wear them simultaneously. Gomersall et al. have reported Spearman $r_S = 0.80$ correlations between ActiGraph GTX3+ and Fitbit One for moderate and vigorous physical activity (N = 29, mean age 39.6, [13]). Participants wore both devices on the hips for two sessions of seven days each. Brewer et al. have found a Pearson $r = 0.69$ correlation between moderate and vigorous physical activity measured by ActiGraph GT3X+ and Fitbit devices (N = 50, university students, [6]). Participants wore the devices on the hip and wrist for seven days. These studies have focused on younger populations and covered a short-term period as well due to the research-grade wearable and the discomfort from wearing both devices.

To our knowledge, no studies quantified the relationships between physical activity PROs and TechROs at different intensities in longitudinal, daily, and free-living conditions for seniors. In our study, we used PROs from the IPAQ validated scale and TechROs from the Fitbit Charge 2 consumer wearable.

Studies have previously assessed physical activity by using the IPAQ. Silsbury et al. have reported that IPAQ has very good reliability and good agreement with accelerometer measures [27]. Van Poppel et al. included IPAQ in the list of scales appropriate for measuring their intended dimension of physical activity [24]. Prior studies in Hungary and Spain, the countries of our research, have also used the IPAQ. In Hungary, Makai et al. used the IPAQ and found differences in physical activity levels by sociodemographic parameters in N = 910 adults [18]. In Spain, Roman-Viñas et al. have found that the IPAQ has good reliability for all its intensities and domains in N = 110 adults [25].

Previous studies measured the accuracy of Fitbit consumer-friendly devices in reporting daily life behaviours of physical activity and sleep. Ferguson et al. have found the Fitbit One, Fitbit Zip, and Withings Pulse to perform the most reliably across physical activity and sleep constructs (N = 21 adults, [9]). Brewer et al. have reported that Fitbit (Charge HR, Charge, Flex, Surge, Zip, Alta) agrees with the ActiWatch GT3X+ accelerometer in assessing active minutes of physical activity in a study run for 7 days (N = 50 university students, [6]). For seniors, Tedesco et al. compared the Fitbit Charge 2 and the Garmin Vivosmart HR+ in free-living environments in a senior cohort (N = 20, age over 65, [31]). They have found that Fitbit had better overall results in step count, energy expenditure, and sleep duration. Paul et al. have also found that Fitbit One and Flex monitor and provide feedback on steps accurately for seniors (N = 32, age over 60, [23]).

The user experience with the wearable of the target group is an essential factor affecting the duration of wear. McMahon et al. have reported that Fitbit One (N = 95, 70+ years old, [20]) is easy to use on the System Usability Scale [7]. Steinert et al. have found that seniors (N = 20, age over 60) graded the Fitbit the highest in usability.

Fitbit provides a well-documented application programming interface (API). The Fitbit API exposes behavioural data for the entire day, including physical activity and sleep.

3. Methods

Our study had three objectives. First, we aimed at quantifying the relationships between PRO intensities and domains of physical activity and TechRO intensities of physical activity. Second, we aimed at identifying stronger relationships beyond the typical questionnaire recall period of several days. Third, we aimed at reporting the quantified relations from data collected while daily life unfolds.

3.1. Study Design

We conducted the study as part of the EU AAL Caregiver and Me (CoME, No. 14-7) research project and software application (2015-2020) for self-management aimed at healthy seniors [2]. The goal of CoME was to reduce the risk of developing dementia [17] by monitoring its risk factors, including physical activity. The institutional review board at the University of Geneva had approved the study in 2016.

Seniors who owned a smartphone, or were willing to receive one for ownership, were invited to a care centre in their city (Lleida, Spain and Budapest, Hungary). A total of 42 individuals (age: 68.78 +/- 6.30; gender: 26 female and 16 male; location: 28 in Spain and 14 in Hungary) agreed to participate from January 2017 to December 2019. They signed written consent. Their identities were pseudonymized.

Upon arrival at the care centre, participants attended an informational workshop about the project. They optionally received a smartphone and a Fitbit Charge 2 consumer wearable as their own (for the study duration and beyond). At the beginning and in subsequent visits throughout 2018 and 2019, the participants answered several questionnaires, including the IPAQ. They were not explicitly informed about when they would fill the questionnaires to avoid any activity pattern change before the visit. Caregivers assisted them throughout the process. Three distinct periods of answers, or *waves*, have resulted: wave 1: mid-2018, wave 2: end-2018 and start-2019, and wave 3: mid-2019.

3.2. Patient-Reported Physical Activity (IPAQ)

The IPAQ long contains 25 questions about the typical duration and frequency of physical activity at *walking*, *moderate*, and *vigorous* intensities, in several domains: *work*, *transport*, *domestic and garden*, and *leisure*. The questions refer to a recall period of one week. The scale provides separate scores of physical activity for each intensity, domain, and overall, derived from the cumulative weekly duration and the energy expenditure. We calculated the IPAQ weekly durations for the intensities and domains (11 variables corresponding to the non-score rows in Table 1). We separately included the scores for the intensities and domains (7 variables). We then added the overall score (1 variable). We obtained a total of 19 PRO variables represented as the rows in Table 1.

We processed the individuals' answers by adhering to the data cleaning, maximum values for excluding outliers, and minimum values for the duration of activity from the IPAQ scoring guideline [14]. The guide does not provide a threshold for converting the *"duration reported as weekly (not daily) to daily into an average daily time"*. For example, if a senior individual reported seven hours of vigorous physical activity per day, the duration would likely reflect one hour per day. We allowed at most 7 hours of physical activity per day in any intensity by dividing all excessive durations by 7 days.

3.3. Technology-Reported Physical Activity (Fitbit)

We assessed the TechRO behaviours of physical activity and sleep. We derived behaviour variables in two amounts, *absolute* and *relative*, with separate semantics. Absolute variables refer to each behaviour separately. Relative variables reflect the difference between a behaviour and the (geometric) mean of all behaviours during the 24 hours of the day. The relative amount was motivated by the interdependence of behaviours during the day [26].

In the absolute amount, we derived the variables directly. For physical activity, we calculated the daily distance (denoted *distance*), energy expenditure (*energy*), step count (*steps*), sedentary duration (*sedentary*), and the duration at three intensities (*light*, *moderate*, and *vigorous*) as reported by Fitbit (7 variables). As Fitbit had not published intensity thresholds, we also derived the cumulative durations in sedentary and light (*sedentary+light*), light and fair (*light+fair*), and fair and vigorous (*fair+vigorous*) intensities (3 variables). We also calculated the total daily active duration (*active*) cumulating the light, fair, and vigorous durations (1 variable). For sleep, we included the entire sleep duration of the day (1 variable). We derived 12 TechRO variables in the absolute amount.

In the relative amount, we derived variables denoting compositional components of physical activity intensities and sleep throughout the day. We derived variables for each component of the centred log-ratio (*CLR*, [1]) transformation. The CLR is a symmetric transformation that does not require a reference component behaviour. We computed the CLRs of two separate compositions: (1) from all physical activity durations (*PAC*) (4 variables) and (2) from all physical activity durations and the sleep duration (*PASC*) (5 variables). We included both relative amounts as the CLRs of a composition are not preserved in sub-compositions [1], but some studies may not be able to monitor sleep. We derived 9 TechRO variables in the absolute amount.

We considered *valid* only those days where the duration covered by wearable monitoring was at least 21 hours. We allowed at most three hours of missing data for device battery charging and handling (15-20 minutes to 2 hours).

Then we derived intervals with fixed durations of 7, 14, 21, 28, 60, 90, and 120 days to balance the number of included days in the analysis with the available intraday monitoring quality. The choice of 7 days for the lower bound was motivated by the need to acquire enough representative data for daily life, the IPAQ recall period of 7 days, and the significant improvements in Fitbit accuracy for active minutes from 7 days onwards [6]. The choice of increasing intervals to the upper bound of 120 days reflected the duration of a wave, a large number of valid days per person (mean 181.1 days), but also the high variance (σ = 179.9 days).

We only included intervals with at least 70% of their days valid, such that both weekdays and weekends were expectedly present in a week; the limit is compatible with previously reported consumer wearable use in seniors [8]. For each interval and variable, we aggregated the mean and geometric mean for the daily absolute and relative amounts, respectively. We included 21 TechRO variables in total, represented as columns in Table 1.

3.4. Statistical Analysis

We aligned in time PROs with TechROs by using the administration date of the PRO answer and the end date of the TechRO measurement interval, with a leeway of at most the interval duration due to scarce exact matches. For each participant, we included only the last alignment in a wave, to discard repeated answers within a few minutes and reduce bias towards overly diligent responders. When we aligned PROs with TechROs of increasing durations, the number of paired observations decreased; we thus required a minimum of 10 observations.

We applied the Spearman rank test measuring the direction and strength of a correlation [29]. We chose this test because the PRO and TechRO assessments were interdependent (they referred to the same participant), not all variables were normally distributed (Shapiro Wilk normality test $p < 0.05$), and the variables had distinct units of measurement (making rank correlation appropriate). The $19 \times 21 = 399$ correlations are depicted in Table 1.

The correlation strength and direction are denoted by a signed real value in the interval $[-1, 1]$, denoted r_S. We considered correlations (r_S) to be *weak* ($r_S \in [0, 0.25]$), *moderate-weak* ($r_S \in (0.25, 0.45]$), *moderate* ($r_S \in (0.45, 0.55]$), *moderate-strong* ($r_S \in (0.55, 0.75]$), or *strong* ($r_S \in (0.75, 1]$) in absence of consensus in the literature. For brevity, we reported only stronger or adjusted correlations. Table 1 depicts the correlations in its cells.

We reported both statistically significant and non-significant correlations. We adjusted the significance following a partial Bonferroni correction by dividing the significance threshold α by 12, the maximum number of interval-agnostic TechRO variables in a given amount, and then by 7, the number of interval durations. The small sample size motivated our choice to balance between no adjustment ($\alpha = 0.05$) and full Bonferroni adjustment ($\alpha = 0.05$ / 19 PRO variables / 21 TechRO variables / 7 interval durations). We placed the significant correlations in three levels of increasing significance: $*$ for $p < 0.05$ (unadjusted), 1 for $p < 0.05 \div 12$ (adjusted), and 2 for $p < 0.05 \div 12 \div 7$ (adjusted). We reported the strength and direction of all correlations regardless of significance. For brevity, we only reported one correlation per TechRO interval duration in the cells of Table 1. If two correlations differed by the TechRO interval duration, we chose the one with a higher level of significance and indicated its duration.

4. Results

We included in further analysis 31 out of 42 initial participants who had filled IPAQ PROs (mean age 70.6 +/- 3.2). The included participants had contributed 53 IPAQ answers (1.71 +/- 0.96 / person) and provided 9836 Fitbit days (317.3 +/- 256.9 / person) from which 5615 were valid days (181.1 +/- 179.2 / person).

As observed in Table 1, the energy expenditure had adjusted moderate-strong correlations with the work moderate and vigorous physical activity. The distance had adjusted moderate-strong correlations with the leisure moderate. The steps had an adjusted moderate-strong correlation with the walking score and the leisure moderate physical activity.

For the absolute physical activity intensities, sedentary duration had negative moderate-strong correlations with the leisure and work walking. Light physical activity duration had adjusted moderate-strong correlations with domestic and leisure moderate physical activity, as well as the score for moderate physical activity. Cumulative light+fair duration had a strong adjusted correlation with the score for moderate physical activity ($r_S = 0.79^1$). Fair duration correlated negatively and moderate-strongly with the moderate physical activity score and the domestic moderate physical activity. Cumulative fair+vigorous duration correlated negatively and moderate-strongly with garden moderate physical activity. The vigorous duration had moderate-strong negative correlations with the domestic and garden score, and garden moderate physical activity. The total active duration correlated moderately-strongly with leisure and moderate scores.

For the PAC relative amount, we report an adjusted negative moderate-strong correlation between the sedentary log-ratio and the work vigorous physical activity as well as unadjusted negative moderate-strong correlations for the transport and overall scores. Light log-ratio had a strong adjusted correlation with the domestic moderate physical activity ($r_S = 0.88^1$) and other moderate-strong correlations in the domestic and garden domain, and the score for

moderate physical activity. Fair log-ratio had a strong correlation with garden vigorous physical activity ($r_S = 0.75^*$) and moderate-strong correlations with moderate and vigorous physical activity. Vigorous log-ratio had a positive correlation with the domestic moderate physical activity and the leisure vigorous physical activity.

For the PASC relative amount (including sleep), sedentary log-ratio had negative and moderate-strong correlations in the work and domestic+garden domains. Light log-ratio had strong correlations with domestic moderate physical activity ($r = 0.84^1$) and the score for moderate physical activity ($r_S = 0.75^*$). Fair log-ratio had strong positive correlations with the work moderate physical activity ($r_S = 0.79^*$) and the vigorous score ($r_S = 0.78^*$). Vigorous log-ratio had a strong positive correlation with the domestic moderate physical activity ($r_S = 0.77^1$), but a negative moderate-strong correlation with the leisure vigorous physical activity.

We report higher correlations in the following objective-subjective pairs as compared to the other pairs where only the TechRO intensity changed. In the absolute amount, we found the strongest correlations between (1) cumulative light+fair duration and moderate physical activity at $r_S = 0.79^1$, (2) light duration and walking at $r_S = 0.64^1$ and moderate physical activity at $r_S = 0.71^1$, (3) fair duration and vigorous physical activity at $r_S = 0.44^*$, (4) cumulative fair+vigorous and vigorous physical activity at $r_S = 0.43^*$, (5) vigorous duration and moderate physical activity at $r_S = 0.46$ (non-significant). In the PAC and PASC relative amounts (1) light CLR correlated the highest with the moderate physical activity at $r_S = 0.68^*$ and $r_S = 0.75^*$, (2) fair CLR correlated the strongest with the vigorous physical activity at $r_S = 0.6^*$ and $r_S = 0.78^*$, and (3) vigorous CLR correlated the strongest again with the moderate physical activity at $r_S = 0.66^*$ and $r_S = 0.63^*$.

Table 1. Rank correlations (cells) between aligned PROs (rows) and TechROs (columns) of physical activity (Spearman r_S)

	Energy	Distance	Steps	Sedentary	Sedentary+Light	Light	Light+Fair	Fair	Fair+Vigorous	Vigorous	Active	Sleep	PAC sedentary	PAC light	PAC fair	PAC vigorous	PASC sedentary	PASC light	PASC fair	PASC vigorous	PASC sleep
Work walking	$0.46^{120,1}$	$0.48^{120,1}$	$0.52^{120,1}$	$0.56^{21,*}$	0.14_	0.33	0.04_	0.15_	0.06_	0.03	$0.45^{14,*}$	0.1_	$0.51^{14,*}$	0.21	0.21_	0.02	$0.56^{14,*}$	0.19	0.34_	0.04	0.05
Work moderate	$0.55^{120,2}$	0.32	$0.4^{120,*}$	$0.49^{21,*}$	0.18_	0.22	0.01	0.14_	0.14_	0.01_	$0.43^{28,*}$	0.12	$0.5^{14,*}$	0.16	$0.74^{120,*}$	0.18	$0.55^{14,*}$	0.26	$0.79^{90,*}$	0.17	0.23
Work vigorous	$0.57^{120,2}$	$0.34^{120,*}$	$0.34^{120,*}$	$0.39^{90,*}$	$0.54^{60,*}$	0.1	0.15_	$0.51^{14,*}$	$0.53^{120,*}$	$0.53^{120,*}$	$0.56^{90,*}$	0.35	$0.61^{14,1}$	0.16_	$0.62^{120,*}$	0.08_	$0.67^{14,1}$	0.01_	$0.55^{7,*}$	0.13_	0.21_
Work score	$0.49^{120,1}$	$0.4^{120,*}$	$0.55^{90,1}$	$0.59^{21,*}$	$0.52^{60,*}$	0.26	0.03_	0.06	0.0	0.11	$0.47^{14,*}$	0.05	$0.48^{28,*}$	0.0	0.38_	0.03	$0.59^{14,*}$	0.12	$0.57^{7,*}$	0.03	0.15
Transport walking	0.26	$0.38^{120,*}$	$0.48^{120,1}$	$0.64^{60,1}$	0.25_	$0.49^{60,*}$	0.32	0.08_	0.44_	0.48_	$0.46^{14,*}$	0.0	$0.5^{14,*}$	$0.56^{60,*}$	0.42_	0.18	$0.46^{14,*}$	$0.63^{90,*}$	0.3_	0.22	0.0_
Transport cycling	$0.4^{21,*}$	$0.44^{60,*}$	$0.47^{28,*}$	$0.49^{21,*}$	$0.38^{14,*}$	0.1	0.09	0.07	0.17	0.18	$0.49^{28,*}$	0.02_	$0.59^{14,*}$	0.34	0.11_	0.05	0.14_	0.46	0.36_	0.03	0.31_
Transport score	0.31	$0.37^{14,*}$	$0.44^{120,*}$	$0.58^{60,*}$	$0.53^{21,*}$	0.34	0.38	0.04_	0.3_	0.37_	$0.56^{90,*}$	$0.55^{21,*}$	$0.62^{14,*}$	$0.63^{120,*}$	0.42_	0.22	0.35_	$0.64^{90,*}$	0.46_	0.25	0.04_
Domestic moderate	$0.49^{120,1}$	0.22	$0.46^{28,*}$	0.25_	0.06	$0.72^{60,1}$	$0.53^{90,*}$	$0.55^{21,*}$	$0.58^{21,*}$	$0.53^{14,*}$	0.06	0.07_	0.52	$0.88^{120,1}$	0.06_	$0.58^{120,*}$	$0.66^{120,*}$	$0.84^{120,1}$	0.05_	$0.77^{28,1}$	$0.62^{28,*}$
Garden moderate	0.04_	0.15	0.19	0.16_	0.11	$0.53^{120,1}$	$0.58^{120,*}$	$0.6^{120,*}$	$0.69^{90,*}$	$0.64^{90,*}$	0.11	0.06_	$0.5^{21,*}$	$0.63^{120,*}$	0.06_	$0.54^{120,*}$	$0.63^{120,*}$	$0.69^{120,*}$	0.15_	$0.58^{120,*}$	0.34
Garden vigorous	$0.44^{120,*}$	$0.44^{120,*}$	$0.46^{120,*}$	$0.42^{90,*}$	$0.45^{90,*}$	0.24	0.11_	0.02	0.1_	0.04	0.0_	$0.43^{120,*}$	0.06	0.06_	$0.75^{120,*}$	0.18	$0.47^{21,*}$	0.14	$0.73^{90,*}$	0.11	0.53
Domestic+garden score	$0.38^{21,*}$	0.17	0.22	0.23_	0.03	$0.43^{120,*}$	$0.53^{90,*}$	0.4_	$0.58^{90,*}$	$0.61^{90,*}$	0.23	0.03	0.44	$0.62^{90,*}$	0.23_	$0.53^{21,*}$	$0.63^{120,*}$	$0.67^{21,1}$	0.31_	$0.66^{120,*}$	0.51
Leisure walking	0.06	$0.48^{28,1}$	$0.51^{60,1}$	$0.66^{60,1}$	0.37_	$0.5^{60,*}$	0.34	0.09_	0.17_	0.23_	$0.54^{120,*}$	$0.45^{28,*}$	0.02_	0.21	0.46_	0.04	0.32_	0.09	0.13_	0.18	0.0_
Leisure moderate	0.09	$0.59^{120,2}$	$0.62^{120,2}$	$0.55^{120,1}$	0.36_	$0.65^{14,1}$	0.52	0.02	0.05_	0.12_	$0.58^{120,*}$	0.0	$0.52^{14,*}$	0.53	0.07_	$0.47^{14,*}$	0.05_	$0.72^{120,*}$	0.07	0.24	0.25_
Leisure vigorous	0.15	0.09	0.04	$0.4^{14,*}$	$0.37^{14,*}$	0.15_	0.01	0.18_	$0.47^{7,*}$	0.42	$0.43^{7,*}$	0.1_	$0.5^{7,*}$	0.27_	$0.59^{60,*}$	$0.7^{28,*}$	0.05_	0.21_	$0.66^{90,*}$	$0.73^{14,*}$	0.17_
Leisure score	0.06	$0.52^{14,1}$	$0.48^{120,1}$	$0.65^{90,1}$	$0.43^{7,*}$	$0.51^{14,*}$	0.4	0.04_	0.03_	0.03_	$0.68^{120,1}$	0.04_	$0.48^{120,*}$	0.3	0.17_	0.16_	$0.46^{7,*}$	0.04_	0.06	0.03_	0.22_
Walking score	0.18	$0.54^{120,1}$	$0.62^{120,2}$	$0.65^{60,1}$	0.25_	$0.64^{60,1}$	0.31	0.13_	0.34_	0.37_	$0.59^{90,*}$	0.07_	$0.51^{14,*}$	0.43	0.46_	0.21	$0.51^{14,*}$	0.23	0.4_	0.25	0.01_
Moderate score	0.09	0.26	$0.39^{120,*}$	$0.41^{14,*}$	0.12	$0.71^{14,1}$	$0.79^{90,1}$	$0.58^{21,*}$	0.52_	$0.56^{21,*}$	$0.62^{14,1}$	0.35_	0.11	$0.68^{90,*}$	0.35_	$0.66^{120,*}$	0.43	$0.75^{120,*}$	0.48_	$0.63^{120,*}$	0.22
Vigorous score	$0.37^{120,*}$	$0.38^{7,*}$	$0.43^{28,*}$	$0.5^{14,1}$	$0.36^{28,*}$	0.0	0.04_	$0.44^{7,*}$	$0.43^{7,*}$	0.46	$0.52^{14,*}$	0.15	$0.5^{14,*}$	0.15_	$0.6^{90,*}$	$0.64^{28,*}$	$0.52^{14,*}$	0.08_	$0.78^{90,*}$	$0.59^{14,*}$	0.04_
Overall score	0.09	$0.39^{120,*}$	$0.48^{120,1}$	$0.41^{14,*}$	0.04	0.35	0.16	0.0_	0.03	0.09_	$0.55^{14,*}$	0.24_	$0.58^{14,*}$	0.52	0.03_	0.1_	$0.49^{14,*}$	0.03	0.17_	0.05_	0.3_

Cells: correlation strength (script), duration and significance (superscript), and direction (subscript), e.g., $0.66^{28,1}$ depicts a (negative) Spearman correlation with $r_S = -0.66$ and $p < 0.05 \div 12$ for an interval with 28 days. Significance: * for $p < 0.05$; 1 for $p < 0.05 \div 12$; 2 for $p < 0.05 \div 12 \div 7$. Colors: orange (weaker correlation) to green (stronger correlation), only for significant correlations.

5. Discussion

The correlations consistently reflected the negative relationship between the objective sedentary duration and the subjective physical activity across all intensities and domains. The sedentary duration had the strongest negative correlations with walking physical activity across all domains and scores, e.g., $r_S = -0.66^1$. The light duration had stronger correlations in the domestic and leisure domains when compared to other domains, e.g., $r_S = 0.72^1$ vs $r_S = 0.49$. Energy expenditure, distance, and steps had mostly moderate correlations in the work and leisure domains, and weaker correlations in the domestic domain, e.g., $r_S = 0.62^2$ vs $r_S = 0.49^1$. The correlations indicate that the seniors engage in physical activity while they are at home or in the garden, but their absolute sedentary time may be unrelated to such activity. Instead, sedentary duration correlated with decreased physical activity in leisure, transport, and work settings, e.g., $r_S = -0.66^1$. Furthermore, energy, distance, and steps did not appear to measure physical

activity at home accurately; energy correlated more with the work domain, e.g., $r_S = 0.57^2$ vs $r_S = 0.44^*$, while steps and distance correlated more with the leisure moderate physical activity, e.g., $r_S = 0.62^2$ vs $r_S = 0.48^1$. This observed difference is consistent with the placement of IPAQ domestic and garden moderate activities in different scoring intensities.

There were stronger correlations across all objective intensities of physical activity for the domestic and garden moderate physical activity, also reflected in the score for moderate physical activity, as compared to other domains, e.g., $r_S = 0.69^*$ vs $r_S = 0.5^*$. This observation indicates that seniors perceive most of the moderate activity to take place around their homes. Objective light duration correlated more with domestic moderate activity, e.g., $r_S = 0.72^1$, while objective fair+vigorous duration correlates more with garden moderate, e.g., $r_S = -0.69^*$. However, the range of objective intensities, e.g., negative correlations, indicates high variability in seniors' descriptions of domestic and garden activities at moderate intensity.

The longitudinal analysis of relative intensities of physical activity leads to stronger correlations between objective and subjective physical activity. In the fair-vigorous intensity spectrum, correlations of absolute intensities are short-term negative, e.g., $r_S = -0.69^*$. However, correlations of relative intensities indicate positive and generally stronger correlations, e.g., $r_S = 0.75^*$. In the PAC relative amount, there are stronger correlations than in the absolute amount for sedentary, light, and fair durations, e.g., $r_S = 0.88^1$ vs $r_S = 0.72^1$. Objectively monitoring seniors longitudinally (up to 120 days) increases the strength of the PRO and TechRO relationships, despite the IPAQ recall period of 7 days.

Measuring sleep in PASC relative amount further strengthened the relations overall from the PAC relative amount in the sedentary to moderate spectrum across all domains, e.g., $r_S = 0.75^*$ vs $r_S = 0.68^*$. Objectively monitoring sleep, in conjunction with physical activity, increased the strength of the physical activity correlations.

Within a small sample size, we report an initial calibration between the definitions of physical activity intensities in TechRO and PRO. In the absolute amount, cumulative light+fair duration and light duration correspond to the moderate physical activity, fair duration corresponds to the vigorous physical activity, and the cumulative fair+vigorous corresponds to the vigorous physical activity. In the relative amounts, the light ratio corresponds to the moderate physical activity, and the fair ratio corresponds to the vigorous physical activity.

Several limitations characterize the study. A first limitation is the presence of multiple answers per individual, but with high variability, for which we only included one answer per wave. A second limitation is a significant decrease in alignments from the original 53 answers; we allowed for a leeway proportional to the interval duration to allow PRO and TechRO alignments that are both (1) short-term, but strict, and (2) longitudinal, but permissive. The study highlights the challenge of retaining individuals (shared by many health studies) that can provide physical activity outcomes through both questionnaire and wearable. A third limitation refers to the simplicity of the chosen variables and the analysis method (correlations with partial adjustment), driven by the reduced sample size.

We expect to employ more advanced techniques and obtain more results within statistical significance as we increase the sample size in further studies aimed at calibrating PROs and TechROs for health outcomes and longitudinal behaviours such as physical activity and sleep in seniors.

Conclusion

We quantified the relationships between physical activity durations reported by the IPAQ questionnaire and the Fitbit wearable in a sample of seniors. Several methodological approaches yielded increasingly stronger relationships across all IPAQ domains and Fitbit intensities, facilitating the calibration of physical activity PROs and TechROs. First, monitoring physical activity longitudinally (beyond the questionnaire recall period). Second, deriving quantifications of physical activity relative to all behaviours throughout the day (compositional). Third, including sleep even in studies targeting physical activity. Our results can inform the design of observational studies that monitor and assess daily life behaviours continuously and longitudinally, and personalized, focused, and effective interventions for senior individuals' targeting physical activity to reduce the risk of chronic disease and improve health and Quality of Life.

Acknowledgements

We thank the projects H2020 WellCO (769765), AAL CoME (14-7), and AAL GUARDIAN (6-120-CP).

References

[1] Aitchison, J., 1982. The statistical analysis of compositional data. J. of the Royal Statistical Society: Series B (Methodological) 44, 139–160.
[2] Ambient Assisted Living, 2020. Caregiver and me. URL: http://come-aal.eu.
[3] Anderson, G., 2011. Responding to the growing cost and prevalence of people with multiple chronic conditions. Johns Hopkins University .
[4] Anderson, K.N., Catt, M., Collerton, J., Davies, K., von Zglinicki, T., Kirkwood, T.B., Jagger, C., 2014. Assessment of sleep and circadian rhythm disorders in the very old: the newcastle 85+ cohort study. Age and Ageing 43, 57–63.
[5] Booth, F.W., Roberts, C.K., Laye, M.J., 2011. Lack of exercise is a major cause of chronic diseases. Comprehensive Physiology 2, 1143–1211.
[6] Brewer, W., Swanson, B.T., Ortiz, A., 2017. Validity of fitbit's active minutes as compared with a research-grade accelerometer and self-reported measures. BMJ Open Sport & Exercise Medicine 3, e000254.
[7] Brooke, J., et al., 1996. Sus-a quick and dirty usability scale. Usability Evaluation in Industry 189, 4–7.
[8] Cooper, C., Gross, A., Brinkman, C., Pope, R., Allen, K., Hastings, S., Bogen, B.E., Goode, A.P., 2018. The impact of wearable motion sensing technology on physical activity in older adults. Experimental Gerontology 112, 9–19.
[9] Ferguson, T., Rowlands, A.V., Olds, T., Maher, C., 2015. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. International Journal of Behavioral Nutrition and Physical Activity 12, 42.
[10] Fitbit, Inc., 2020. Fitbit official website. URL: https://www.fitbit.com.
[11] Garriguet, D., et al., 2015. Comparison of physical activity adult questionnaire results with accelerometer data. Health Reports 26, 11.
[12] General, S., 1979. The surgeon general's report on health promotion and disease prevention. DHEW (PHS) Publication .
[13] Gomersall, S.R., et al., 2016. Estimating physical activity and sedentary behavior in a free-living context: a pragmatic comparison of consumer-based activity trackers and actigraph accelerometry. Journal of Medical Internet Research 18, e239.
[14] Hallal, P.C., Victora, C.G., 2004. Reliability and validity of the international physical activity questionnaire (ipaq). Medicine & Science in Sports & Exercise 36, 556.
[15] Inkin, 2020. Wearables: the comprehensive list of wearables on the market. URL: https://www.inkin.com/wearables/.
[16] International Data Corporation, 2020. IDC forecasts sustained double-digit growth for wearable devices led by steady adoption of smart-watches. URL: https://www.idc.com/getdoc.jsp?containerId=prUS44553518.
[17] Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S.G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., et al., 2017. Dementia prevention, intervention, and care. The Lancet 390, 2673–2734.
[18] Makai, A., Füge, K., Breitenbach, Z., Figler, M., 2016. Physical activity levels of adults in two counties of hungary in view of sociodemographic inequalities. Health Problems of Civilization 10, 38–46.
[19] Mayo, N.E., Figueiredo, S., Ahmed, S., Bartlett, S.J., 2017. Montreal accord on patient-reported outcomes (pros) use series–paper 2: terminology proposed to measure what matters in health. Journal of Clinical Epidemiology 89, 119–124.
[20] McMahon, S.K., Lewis, B., Oakes, M., Guan, W., Wyman, J.F., Rothman, A.J., 2016. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR mHealth and uHealth 4, e35.
[21] Naghavi, M., Abajobir, A.A., Abbafati, C., et al., 2017. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the global burden of disease study 2016. The Lancet 390, 1151–1210.
[22] National Cancer Institute, 2020. Validation of Physical Activity. URL: https://epi.grants.cancer.gov/paq/validation.html.
[23] Paul, S.S., Tiedemann, A., Hassett, L.M., Ramsay, E., Kirkham, C., Chagpar, S., Sherrington, C., 2015. Validity of the fitbit activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport & Exercise Medicine 1.
[24] van Poppel, M.N., et al., 2010. Physical activity questionnaires for adults. Sports Medicine 40, 565–600.
[25] Roman-Viñas, B., Serra-Majem, L., Hagströmer, M., Ribas-Barba, L., Sjöström, M., Segura-Cardona, R., 2010. International physical activity questionnaire: reliability and validity in a spanish population. European Journal of Sport Science 10, 297–304.
[26] Rosenberger, M.E., Fulton, J.E., Buman, M.P., Troiano, R.P., Grandner, M.A., Buchner, D.M., Haskell, W.L., 2019. The 24-hour activity cycle: a new paradigm for physical activity. Medicine & Science in Sports & Exercise 51, 454.
[27] Silsbury, Z., Goldsmith, R., Rushton, A., 2015. Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. BMJ Open 5, e008430.
[28] Sims, J., Smith, F., Duffy, A., Hilton, S., 1999. The vagaries of self-reports of physical activity: a problem revisited and addressed in a study of exercise promotion in the over 65s in general practice. Family Practice 16, 152–157.
[29] Spearman, C., 1961. The proof and measurement of association between two things. .
[30] Sylvia, L.G., et al., 2014. Practical guide to measuring physical activity. Journal of the Academy of Nutrition and Dietetics 114, 199–208.
[31] Tedesco, S., Sica, M., Ancillao, A., Timmons, S., Barton, J., O'Flynn, B., 2019. Validity evaluation of the fitbit charge2 and the garmin vivosmart hr+ in free-living environments in an older adult cohort. JMIR mHealth and uHealth 7, e13084.
[32] US FDA, 2020. Digital Health Software Precertification (Pre-Cert) Program. URL: https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program.
[33] Van Den Berg, J.F., Van Rooij, F.J., Vos, H., Tulen, J.H., Hofman, A., Miedema, H.M., Neven, A.K., Tiemeier, H., 2008. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. Journal of Sleep Research 17, 295–302.
[34] Wanner, M., Probst-Hensch, N., Kriemler, S., Meier, F., Autenrieth, C., Martin, B.W., 2016. Validation of the long international physical activity questionnaire: influence of age and language region. Preventive Medicine Reports 3, 250–256.
[35] WHO, 2020. Global recommendations on physical activity for health. URL: https://www.who.int/dietphysicalactivity/global-PA-recs-2010.pdf.
[36] WHO regional office for Europe, 2020. Noncommunicable diseases. URL: http://www.euro.who.int/en/health-topics/noncommunicable-diseases.

Publication 8: Using Consumer-Friendly Wearables to Associate Patient-Reported Quality of Life and Tech-Reported Physical Activity and Sleep in Healthy Seniors

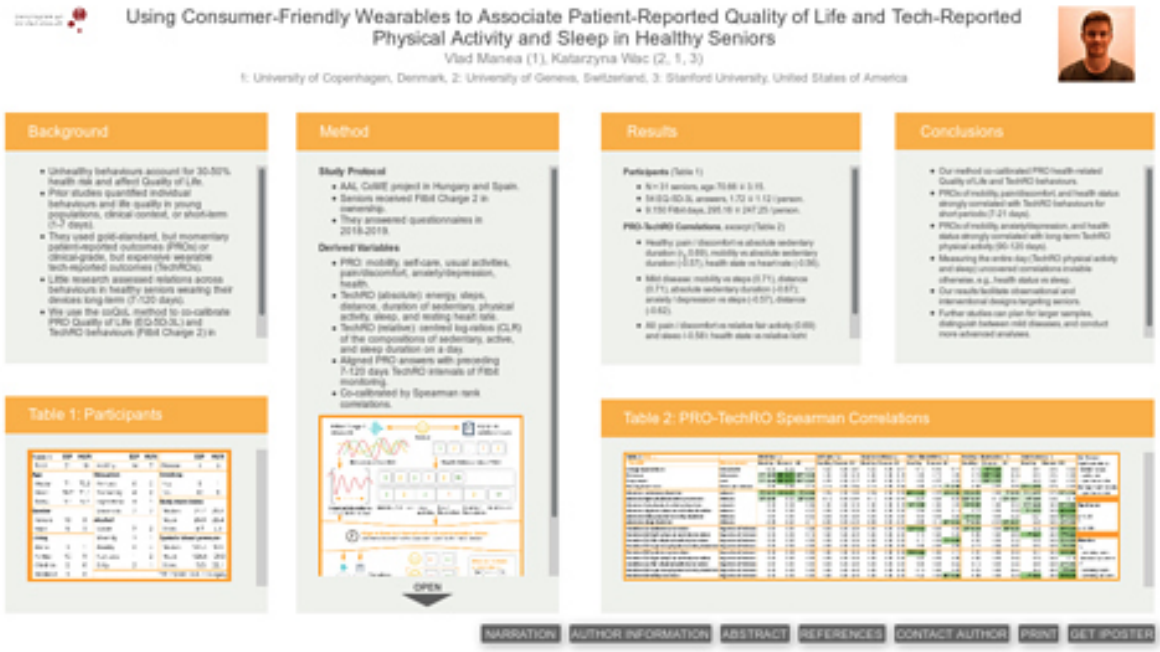
Vlad Manea, Katarzyna Wac. *Using Consumer-Friendly Wearables to Associate Patient-Reported Quality of Life and Tech-Reported Physical Activity and Sleep in Healthy Seniors*. Poster at the Conference of the International Society for Quality of Life Research (ISOQOL 2020). DOI: https://doi.org/10.1007/s11136-020-02626-y. [57]

Abstract (generated) Behaviours account for 50% health risk and affect life quality later in life. Numerous studies quantified the relationships between isolated behaviours and life quality in clinical cases, short-term, using momentary reported outcomes or expensive wearables. However, little research studied relations across multiple behaviours in healthy seniors wearing their own devices long-term (7-120 days). Methods 42 seniors in Spain and Hungary (aged 68.78 +/- 6.30) patient-reported Quality of Life (EQ-5D-3L) and tech-reported daily life behaviours (Fitbit Charge 2). We align answers to intervals (7-120 days) by administration date and end date, within a leeway proportional to the interval duration. We derive patient-reported variables and tech-reported variables (energy, steps, distance, duration of sedentary, activity, sleep, and resting heart rate) in absolute and, where relevant, relative (compositional) quantities. We quantify Spearman associations at alpha = 0.05. N = 31 participants (aged 70.66 $\pm$ 3.15: 21 in Spain and 10 in Hungary) provided 54 EQ-5D-3L answers (1.72 $\pm$ 1.12 / person) and 9.150 Fitbit days (295.16 $\pm$ 247.25 / person). 10 participants reported mild disease. In all participants, distance and steps associated with mobility ($r = 0.71$), $p < 0.005$. Sleep duration inversely associated with anxiety (-0.57) and pain (-0.52): vigorous duration associated with health state (0.70): relative light activity associated with health state (0.63), $p < 0.005$. In healthy participants, absolute sedentary duration associated with a lack of mobility (0.57) and pain (0.69), and a high resting heart rate associated with poor health (0.56), $p < 0.005$. Relative sedentary duration associated with pain (0.62) and lack of anxiety (0.54) while light relative duration associated with health state (0.64). Relative sleep duration associated with health state (0.65). In sick participants, distance and steps associated with mobility (0.71) and lack of anxiety (-0.57), stronger, less significant, and only over longer periods than absolute quantities. Conclusions Our method is feasible in associating behaviours and mobility, pain, and health status for short periods (7-21 days) in a small sample of healthy participants. Monitoring physical activity log-term (90-120 days) helped better assess mobility,

anxiety, and health state in sick seniors. Our results provide insights for designs targeting interventions for seniors.

Keywords observational study, healthy senior, quality of life, statistical correlation, questionnaire, EQ-5D-3L, consumer-friendly wearable, Fitbit.

Using Consumer-Friendly Wearables to Associate Patient-Reported Quality of Life and Tech-Reported Physical Activity and Sleep in Healthy Seniors



Vlad Manea (1), Katarzyna Wac (2, 1, 3)

1: University of Copenhagen, Denmark, 2: University of Geneva, Switzerland, 3: Stanford University, United States of America

PRESENTED AT:

VIRTUAL ISOQOL 2020

2020 THE FUTURE IS NOW:
A Vision for the Future of Outcomes Measurement and Quality of Life Research

BACKGROUND

- Unhealthy behaviours account for 30-50% health risk and affect Quality of Life.
- Prior studies quantified individual behaviours and life quality in young populations, clinical context, or short-term (1-7 days).
- They used gold-standard, but momentary patient-reported outcomes (PROs) or clinical-grade, but expensive wearable tech-reported outcomes (TechROs).
- Little research assessed relations across behaviours in healthy seniors wearing their devices long-term (7-120 days).
- We use the coQoL method to co-calibrate PRO Quality of Life (EQ-5D-3L) and TechRO behaviours (Fitbit Charge 2) in seniors.

TABLE 1: PARTICIPANTS

Table 1	ESP	HUN		ESP	HUN		ESP	HUN
Total	21	10	Healthy	14	7	Disease	7	3
Age			**Education**			**Smoking**		
Median	71	70.5	Primary	6	0	Yes	5	1
Mean	70.7	71.1	Secondary	4	2	No	16	9
Stdev.	3.1	8.7	Highschool	4	1	**Body mass index**		
Gender			University	7	7	Median	24.7	26.8
Female	13	5	**Alcohol**			Mean	26.0	28.4
Male	8	5	Never	7	2	Stdev.	4.7	4.1
Living			Monthly	3	5	**Systolic blood pressure**		
Alone	9	1	Weekly	8	0	Median	125.5	120
Partner	10	9	Few days	1	2	Mean	124.4	129.5
Children	2	0	Daily	2	1	Stdev.	15.5	22.1
Nephews	0	0				ESP = Spain; HUN = Hungary		

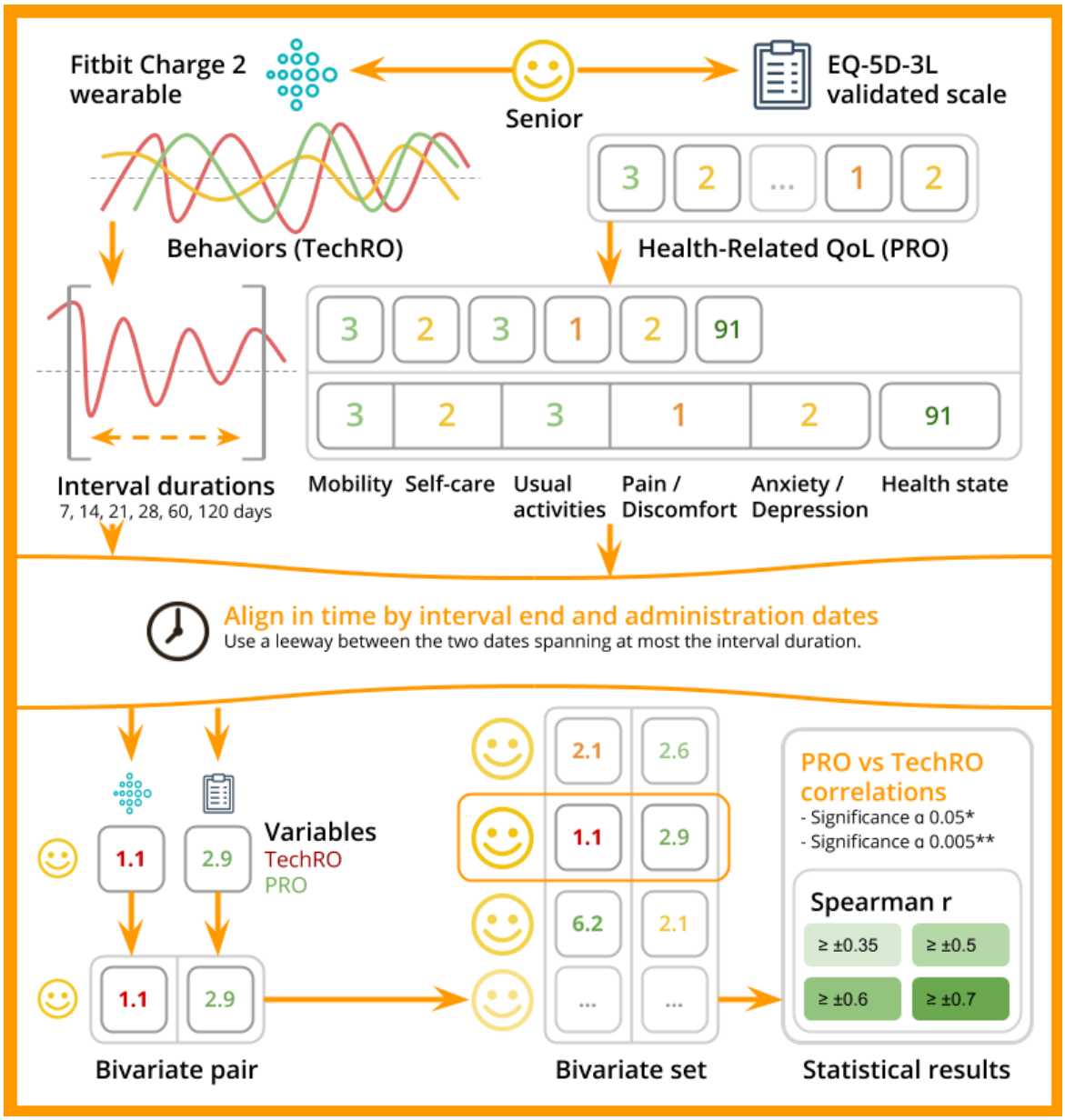
METHOD

Study Protocol

- AAL CoME project in Hungary and Spain.
- Seniors received Fitbit Charge 2 in ownership.
- They answered questionnaires in 2018-2019.

Derived Variables

- PRO: mobility, self-care, usual activities, pain/discomfort, anxiety/depression, health.
- TechRO (absolute): energy, steps, distance, duration of sedentary, physical activity, sleep, and resting heart rate.
- TechRO (relative): centred log-ratios (CLR) of the compositions of sedentary, active, and sleep duration on a day.
- Aligned PRO answers with preceding 7-120 days TechRO intervals of Fitbit monitoring.
- Co-calibrated by Spearman rank correlations.



RESULTS

Participants (Table 1)

- N = 31 seniors, age 70.66 ∓ 3.15.
- 54 EQ-5D-3L answers, 1.72 ∓ 1.12 / person.
- 9.150 Fitbit days, 295.16 ∓ 247.25 / person.

PRO-TechRO Correlations, excerpt (Table 2)

- Healthy: pain / discomfort vs absolute sedentary duration (r_S 0.69), mobility vs absolute sedentary duration (-0.57), health state vs heart rate (-0.56).
- Mild disease: mobility vs steps (0.71), distance (0.71), absolute sedentary duration (-0.67); anxiety / depression vs steps (-0.57), distance (-0.62).
- All: pain / discomfort vs relative fair activity (0.69) and sleep (-0.58); health state vs relative light activity duration (0.63) and sleep duration (0.73).

CONCLUSIONS

- Our method co-calibrated PRO health-related Quality of Life and TechRO behaviours.
- PROs of mobility, pain/discomfort, and health status strongly correlated with TechRO behaviours for short periods (7-21 days).
- PROs of mobility, anxiety/depression, and health status strongly correlated with long-term TechRO physical activity (90-120 days).
- Measuring the entire day (TechRO physical activity and sleep) uncovered correlations invisible otherwise, e.g., health status vs sleep.
- Our results facilitate observational and interventional designs targeting seniors.
- Further studies can plan for larger samples, distinguish between mild diseases, and conduct more advanced analyses.

TABLE 2: PRO-TECHRO SPEARMAN CORRELATIONS

Table 2 \ PRO →		Mobility (↘)			Self-care (↘)			Usual activities (↘)			Pain / discomfort (↗)			Anxiety / depression (↗)			Health state (↗)			Cell Format
↕ TechRO	Measurement	Healthy	Disease	All	Healthy	Disease	All	Healthy	Disease	All	Healthy	Disease	All	Healthy	Disease	All	Healthy	Disease	All	Significant results:
Energy expenditure	kilocalorie	-0.25	-0.25	-0.21	0.25	0.00	0.17	0.25	0.00	0.17	-0.13	-0.45	-0.14	0.15	14* -0.64	0.02	0.11	0.16	0.03	- duration (days)
Distance	kilometer	21* -0.39	90* -0.71	28** -0.45	-0.30	0.00	-0.28	-0.30	0.00	-0.28	0.05	-0.19	0.05	-0.18	120* -0.62	-0.26	-0.14	0.15	0.00	- significance
Step count	unit	21* -0.40	90* -0.71	90** -0.47	-0.35	0.00	-0.27	-0.35	0.00	-0.27	0.02	-0.23	-0.01	-0.12	120* -0.57	-0.21	-0.09	0.41	-0.02	- signed correlation
Resting heart rate	beats per minute	0.35	0.06	0.16	0.16	0.00	0.14	0.16	0.00	0.14	0.00	-0.10	7* -0.38	-0.08	-0.39	-0.19	21** -0.56	-0.20	-0.20	Non-significant results:
Absolute sedentary duration	minute	7** 0.57	21* 0.67	7** 0.56	0.24	0.00	0.18	0.24	0.00	0.18	60** 0.69	0.03	60* 0.50	120* 0.53	0.00	7* 0.35	21* -0.37	-0.23	21* -0.34	- signed correlation
Absolute light physical activity duration	minute	14* -0.40	-0.22	60* -0.49	-0.12	0.00	-0.14	-0.12	0.00	-0.14	-0.27	0.20	-0.25	120* -0.47	-0.12	90* -0.41	14* 0.44	0.12	0.14	
Absolute fair physical activity duration	minute	-0.30	0.00	-0.06	-0.30	0.00	-0.28	-0.30	0.00	-0.28	14* -0.55	0.00	0.03	21* 0.52	0.00	-0.03	21* -0.46	0.00	21* -0.50	**Significance**
Absolute vigorous physical activity duration	minute	0.00	0.00	-0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.11	0.15	0.00	-0.19	7* -0.50	0.00	14* -0.70	*:
Absolute daily physical activity duration	minute	0.00	0.00	-0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	21* 0.54	0.00	-0.30	-0.21	0.00	14* -0.56	p < 0.05
Absolute sleep duration	minute	-0.38	-0.26	-0.11	-0.38	0.00	-0.30	-0.38	0.00	-0.30	-0.45	0.12	60* -0.52	7* -0.43	-0.03	60* -0.57	0.40	0.06	28* 0.44	**:
Relative CLR sedentary duration	log ratio of minute	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.00	0.00	120* 0.62	0.00	0.08	21* -0.54	0.00	90* 0.59	0.32	0.00	14* 0.64	p < 0.005
Relative CLR light physical activity duration	log ratio of minute	0.00	0.00	-0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	-0.26	-0.30	0.00	-0.15	7* 0.64	0.00	7** 0.63	
Relative CLR fair physical activity duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.00	21* 0.69	-0.02	0.00	0.23	-0.15	0.00	21* 0.55	**Direction**
Relative CLR vigorous physical activity duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.26	-0.19	0.00	0.07	0.52	0.00	7* 0.50	↘:
Relative CLR sedentary duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	120* 0.64	0.00	-0.13	-0.01	0.00	0.30	0.22	0.00	14* 0.62	- increasing score
Relative CLR light physical activity duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.00	-0.30	0.00	0.00	-0.03	0.33	0.00	0.14	- decreasing outcome
Relative CLR fair physical activity duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.13	0.00	0.42	0.00	0.00	21* 0.49	↗:
Relative CLR vigorous physical activity duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.19	0.00	-0.41	-0.25	0.00	-0.25	0.53	0.00	7* 0.54	- increasing score
Relative CLR sleep duration	log ratio of minute	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.22	0.00	90* -0.58	-0.38	0.00	0.30	7* 0.65	0.00	14* 0.73	- increasing outcome

Because of maintenance we will within a few minutes restart our server. We will be back in a moment.

Sorry for the inconvenience!

AUTHOR INFORMATION

Vlad Manea

University of Copenhagen

manea@di.ku.dk

Katarzyna Wac

University of Geneva

University of Copenhagen

Stanford University

katarzyna.wac@unige.ch

Quality of Life Technologies Lab

qualityoflifetechnologies.com (https://qualityoflifetechnologies.com)

ABSTRACT

Aims Behaviours account for 50% health risk and affect life quality later in life. Numerous studies quantified the relationships between isolated behaviours and life quality in clinical cases, short-term, using momentary reported outcomes or expensive wearables. However, little research studied relations across multiple behaviours in healthy seniors wearing their own devices long-term (7-120 days). Methods 42 seniors in Spain and Hungary (aged 68.78 +/- 6.30) patient-reported Quality of Life (EQ-5D-3L) and tech-reported daily life behaviours (Fitbit Charge 2). We align answers to intervals (7-120 days) by administration date and end date, within a leeway proportional to the interval. duration We derive patient-reported variables and tech-reported variables (energy, steps, distance, duration of sedentary, activity, sleep, and resting heart rate) in absolute and, where relevant, relative (compositional) quantities. We quantify Spearman associations at alpha = 0.05. Results N = 31 participants (aged 70.66 +/- 3.15: 21 in Spain and 10 in Hungary) provided 54 EQ-5D-3L answers (1.72 +/- 1.12 / person) and 9.150 Fitbit days (295.16 +/- 247.25 / person). 10 participants reported mild disease. In all participants, distance and steps associated with mobility ($r = 0.71$), $p < 0.005$. Sleep duration inversely associated with anxiety (-0.57) and pain (-0.52): vigorous duration associated with health state (0.70): relative light activity associated with health state (0.63), $p < 0.005$. In healthy participants, absolute sedentary duration associated with a lack of mobility (0.57) and pain (0.69), and a high resting heart rate associated with poor health (0.56), $p < 0.005$. Relative sedentary duration associated with pain (0.62) and lack of anxiety (0.54) while light relative duration associated with health state (0.64). Relative sleep duration associated with health state (0.65). In sick participants, distance and steps associated with mobility (0.71) and lack of anxiety (-0.57), stronger, less significant, and only over longer periods than absolute quantities. Conclusions Our method is feasible in associating behaviours and mobility, pain, and health status for short periods (7-21 days) in a small sample of healthy participants. Monitoring physical activity log-term (90-120 days) helped better assess mobility, anxiety, and health state in sick seniors. Our results provide insights for designs targeting interventions for seniors.

REFERENCES

1. People, H. (1979). The Surgeon General's report on health promotion and disease prevention. Washington, DC: US Dept of Health, Education, and Welfare, 177.

2. Yurkiewicz, I. R., Simon, P., Liedtke, M., Dahl, G., & Dunn, T. (2018). Effect of Fitbit and iPad wearable technology in health-related quality of life in adolescent and young adult cancer patients. Journal of adolescent and young adult oncology, 7(5), 579-583.

3. Tao, K., Liu, W., Xiong, S., Ken, L., Zeng, N., Peng, Q., ... & Li, X. (2019). Associations between self-determined motivation, accelerometer-determined physical activity, and quality of life in Chinese college students. International journal of environmental research and public health, 16(16), 2941.

4. van Uem, J. M., Isaacs, T., Lewin, A., Bresolin, E., Salkovic, D., Espay, A. J., ... & Maetzler, W. (2016). A viewpoint on wearable technology-enabled measurement of wellbeing and health-related quality of life in Parkinson's disease. Journal of Parkinson's disease, 6(2), 279-287.

5. Mayo, N. E., Figueiredo, S., Ahmed, S., & Bartlett, S. J. (2017). Montréal Accord on Patient-Reported Outcomes (PROs) use series–Paper 2: Terminology proposed to measure what matters in health. Journal of clinical epidemiology, 89, 119-124.

6. Wac, K. (2018). From quantified self to quality of life. In Digital Health (pp. 83-108). Springer, Cham.

7. Spearman, C. (1961). The proof and measurement of association between two things.

8. Fitbit (2020). fitbit.com (https://fitbit.com)

9. EuroQoL (2020). EQ-5D-3L. euroqol.org/eq-5d-instruments/eq-5d-3l-about (https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/)

10. AAL Caregiver & Me (2020). come-aal.eu (http://come-aal.eu)

11. Aitchison, J. (1982). The statistical analysis of compositional data. Journal of the Royal Statistical Society: Series B (Methodological), 44(2), 139-160.

Publication 9: Energy and Fatigue. Classification and Assessment of Energy and Fatigue using Subjective, Objective, and Mixed Methods towards Health and Quality of Life

Natalie Solomon, Vlad Manea. *Energy and Fatigue: Classification and Assessment of Energy and Fatigue using Subjective, Objective, and Mixed Methods towards Health and Quality of Life* (accepted). Book chapter in: Katarzyna Wac, Sharon Wulfovich (eds.), *Quantifying Quality of Life: Incorporating Daily Life into Medicine*, Health Informatics, Springer, Cham. 30p. [58]

Abstract Energy and fatigue carry important implications for vitality and overall quality of life. Lacking energy and experiencing fatigue can be both burdensome as well as adaptive. This chapter first classifies energy and fatigue and then reviews their measurement. This chapter closes with opportunities for future directions. Energy and fatigue are present under varying conditions including in daily performance, during and after acute physical or mental strain (capacity), and in the context of chronic conditions. Energy and fatigue have been measured both subjectively and objectively. Subjective outcomes can be derived from self-reported scales and prompts; objective outcomes derived from performance and capacity tasks and technology-reported physiological, biological, and behavioural markers. The scales and tasks employed to measure energy have been traditionally validated but may lack daily life context and ecological validity. Prompts and behavioural monitoring methods are emerging as promising alternatives. Energy and fatigue have also been routinely monitored for specific diseases and occupations. However, fewer studies monitor healthy individuals through consumer technology in daily life contexts. More research is needed for an objective, unobtrusive, longitudinal, and contextual measurement of energy and fatigue in the healthy general population, in service of improving health, wellbeing, and quality of life.

Keywords energy, fatigue, subjective methods, objective methods, mixed methods, quality of life.

UNIVERSITÉ DE GENÈVE

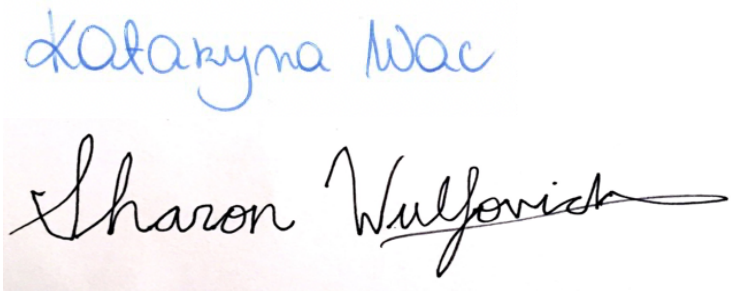
To whom it may concern,

The chapter *Quantifying Energy and Fatigue life* by Natalie Solomon and Vlad Manea and has been accepted for publication in the forthcoming edited book *Quantifying Quality of Life: Incorporating Daily Life into Medicine* published by Springer Nature Switzerland AG (Under the imprint Springer) to be published in 2021.

If you have any questions, please feel free to reach me at

wulfovichsharon@gmail.com

Sincerely,
Prof. Katarzyna Wac and Sharon Wulfovich
Quality of Life Technologies Lab
University of Copenhagen (DK) & University of Geneva (CH) & UC San Diego School of Medicine
https://www.qualityoflifetechnologies.com/



Quantifying Energy and Fatigue

Classification and Assessment of Energy and Fatigue using Subjective, Objective, and Mixed Methods towards Health and Quality of Life

Natalie Solomon, Psy.D., PGSP Stanford Psy.D. Consortium

Vlad Manea, M.Sc., University of Copenhagen

Abstract

Energy and fatigue carry important implications for vitality and overall quality of life. Lacking energy and experiencing fatigue can be both burdensome as well as adaptive. This chapter first classifies energy and fatigue and then reviews their measurement. This chapter closes with opportunities for future directions.

Energy and fatigue are present under varying conditions including in daily performance, during and after acute physical or mental strain (capacity), and in the context of chronic conditions. Energy and fatigue have been measured both subjectively and objectively. Subjective outcomes can be derived from self-reported scales and prompts; objective outcomes derived from performance and capacity tasks and technology-reported physiological, biological, and behavioural markers. The scales and tasks employed to measure energy have been traditionally validated but may lack daily life context and ecological validity. Prompts and behavioural monitoring methods are emerging as promising alternatives.

Energy and fatigue have also been routinely monitored for specific diseases and occupations. However, fewer studies monitor healthy individuals through consumer technology in daily life contexts. More research is needed for an objective, unobtrusive, longitudinal, and contextual measurement of energy and fatigue in the healthy general population, in service of improving health, wellbeing, and quality of life.

Keywords

Energy, fatigue, vitality, fatigue taxonomy, pathological fatigue, non-pathological fatigue, physical fatigue, fatigability, mental fatigue, burnout, fatigue scale, fatigue task, subjective measurement, objective measurement, mixed methods, validated scale, ecological momentary assessment, research sensor, consumer wearable, measurement property spectrum, longitudinal behaviour monitoring.

Introduction

There are many ways to conceptualize “Energy” and “Fatigue” in the context of the WHO Quality of Life domain [1]. Energy and fatigue may be interrelated but may also be considered orthogonal. Low energy can be characterized by fatigue, lack of motivation, and lack of interest, while states of excessive energy can reach pathological levels that include disrupted sleep, restlessness and agitation, or even mania [2]. Although lacking energy can be burdensome and uncomfortable, it is simultaneously an adaptive symptom that is perceived as a need to rest or slow down [3]. Given that energy is a valuable resource, efficient spending and conservation of energy may result in the greatest chances of vitality and even survival [4]. Fatigue is both a normative experience as well as associated with many chronic illnesses and psychiatric disorders. Fatigue can be characterized by subjective feelings of “tiredness” and “lack of energy” [5] and can serve as a signal to prevent strain, damage, and injury [6].

In this chapter, energy refers to the strength and vitality required for sustained physical or mental activity. Lack of energy or fatigue is used to describe the subjective sensation (*perceived fatigue*) as well as the objective and quantifiable change in performance (*fatigability*) [7]. Fatigue can be classified as pathological or non-pathological. Pathological fatigue can be described as an overwhelming sense of tiredness at rest, exhaustion with activity, lack of energy that precludes daily tasks, or loss of vigour [7]. In healthy adults, non-pathological fatigue is predictable and does not interfere with usual daily activities. Non-pathological fatigue is typically brought about by prolonged exertion and diminishes with rest [8]. In addition to pathological and non-pathological fatigue, fatigue may also be subdivided as either physical or mental (cognitive/psychiatric) and further subdivided as primary (neurological) or secondary (non-neurological) [7]–[10]. Furthermore, performance refers to an individual’s functioning in their daily environment while capacity refers to the maximal or optimized level of functioning.

Preliminary studies were conducted on energy and fatigue during the First World War when researchers investigated the impact of fatigue on efficiency and productivity of the industrial workforce [11]. This “occupational fatigue” continues to be a focus of research attention, especially in vocations and occupations in which fatigue carries serious implications. Traditionally, energy and fatigue have been assessed using qualitative, self-reported outcomes [12] and can be obtained from a number of validated scales [13], [14]. Most clinical fatigue studies use self-report measures that can broadly be classified as measuring perceptions of fatigue [8]. Despite the numerous scales that measure fatigue, there is no agreed-upon standard of which to compare subjective reports of fatigue [15], [16].

The use of technology to monitor and manage energy and fatigue has been investigated in order to help healthy individuals continue to live healthily [3], [6], [17], assist individuals with health issues [18]–[20], and address vocational or occupational fatigue to improve personal and workplace safety [21]–[23]. The monitoring of energy and fatigue helps individuals adapt their effort in recreational (e.g., amateur sport, exercise) and occupational (e.g., drivers, pilots, police, professional athletes, shift workers) settings to prevent negative effects (e.g., burnout,

exhaustion, accidents, injury) and maintain quality of life [24], [25]. Further, energy and fatigue research is needed to examine their connection to underlying or potential health conditions as well as interventional studies to validate the operationalization of energy and fatigue monitoring in daily life.

In this chapter, we will classify energy and fatigue and present their measurement. The chapter is structured as methods of our work, classification of energy and fatigue (pathological as well as non-pathological), measurement and assessment of energy and fatigue, discussion of results, and conclusive remarks.

Methods

We conducted a scoping review of the existing literature between 2010 and 2020 in Google Scholar on the technology-enabled assessment of energy and fatigue. Search terms related to energy and fatigue (e.g., “fatigability”, “tiredness”) were coupled with terms pertaining to each of the following domains: (1) the population under study (e.g., “athlete”, “driver”), (2) the health outcomes (e.g., “circulation”, “dementia”, “heart”), and (3) the measurement (e.g., “accelerometer”, “electrocardiogram”, “wearable”). One example search phrase was “galvanic energy fatigue tiredness vitality”. We also reviewed the relevant references of the identified literature. Table 1 reviews the domains, search terms (selective), and the rationale for choosing the domain.

Table 1. Domains of energy and fatigue literature review

Domain	**Inclusion**	**Rationale**	**Search terms (selective)**
Energy / Fatigue	Mandatory	Energy and fatigue are often proxied by synonyms or antonyms.	energy, fatigue, fatigability, tiredness, vitality
Population	Optional	Papers assessing energy and fatigue in healthy individuals often focus on a specific segment of the general population. For instance, two areas of focus are athletics and occupational fatigue.	athlete, driver, performance, pilot, police, shift, sport, worker, employee
Health outcomes	Optional	Health outcomes are often delineated by specific elements of human physiology or pathology: organs, systems, processes, and diseases. In addition, such elements can be further delineated by the	cancer, cardiovascular, circulation, dementia, heart, kidney, mental, pulmonary, respiration

		population segment under study.	
Measurement	Optional	Methodological measurements using technology can be described by the procedure, device, sensor, process, or result.	accelerometer, app, application, camera, band, ecological momentary assessment, performance, capacity, electrocardiogram, electrooculogram, experience sampling method, Fitbit, galvanic, mobile, sensor, smart band, smartphone, smartwatch, vision, watch, wearable

Results

We found 40 reviews on energy and fatigue pertaining to the domains and 60 studies assessing fatigue by using technology. The search results included in this review either (1) reviewed energy and fatigue assessment for a specific population and/or health outcome, (2) provided evidence for the use of measurement to monitor or manage energy or fatigue, or (3) discussed human factors of technology towards monitoring energy and fatigue. The taxonomy of fatigue resulting from our literature review is depicted in Figure 1.

Energy and Fatigue Classification

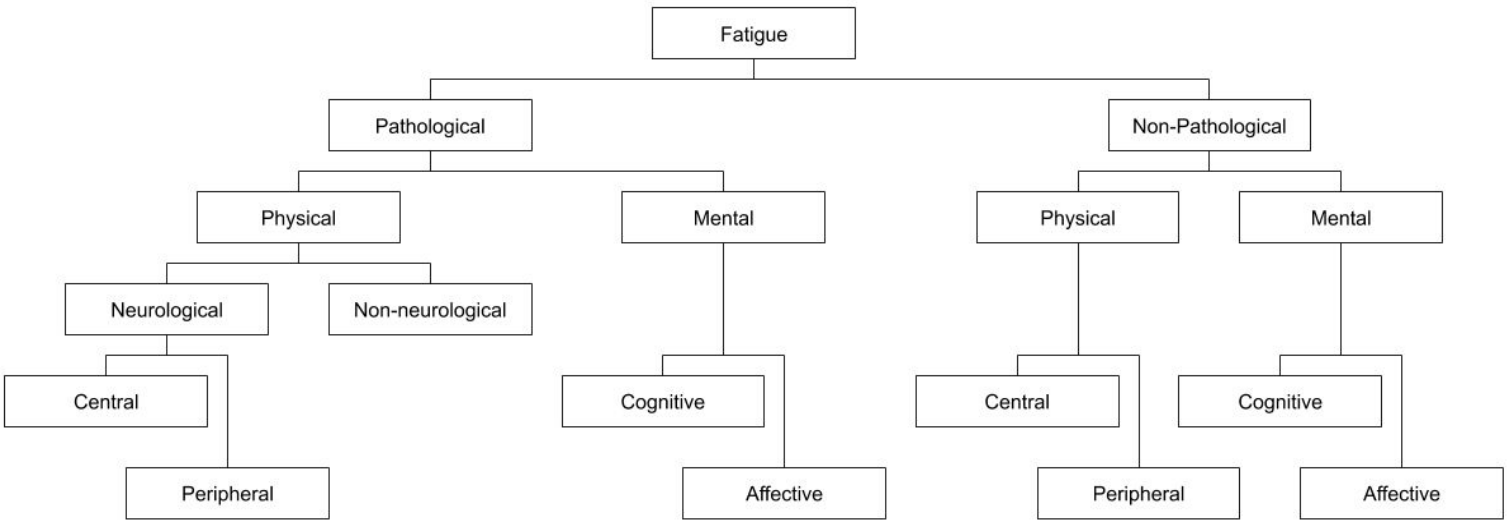


Figure 1. Taxonomy of fatigue with pathological and non-pathological types.

Modified from Chaudhuri & Behan [9], Finsterer & Mahjoub [7], Glaus [26], Kluger, Krupp, & Enoka [8], and Mollayeva et al. [10].

Pathological Fatigue

Pathological fatigue is prolonged or chronic (>6 months), can be highly debilitating, and is much less common than normal fatigue [27]. Pathological fatigue may be best understood as an amplified sense of normal (non-pathological) fatigue that can be induced by changes in one or more variables regulating work output [9]. For instance, a healthy individual may experience fatigue during or after exercising, but the same individual may perceive even more fatigue when exercising during an infectious disease [7]. Diseased individuals describe fatigue as an overwhelming sense of tiredness at rest, exhaustion with activity, loss of vigour, or lack of energy that precludes daily tasks, inertia or lack of endurance [28]. Pathological fatigue may be classified as physical or mental and is associated with multiple illnesses.

Physical Fatigue

Pathological physical fatigue includes neurological and non-neurological fatigue.

Neurological Fatigue

Neurological fatigue suggests that the physical expression of fatigue is mediated by central and peripheral mechanisms [27]. Therefore, neurological fatigue may be further classified as central or peripheral [9].

Central fatigue is generated at sites proximal to the peripheral nerves and referred to as a progressive decline in the ability to activate muscles voluntarily [29]. Central fatigue is due to impaired muscle performance that arises from the central nervous system [28]. A feeling of constant exhaustion is a characteristic of central fatigue [9]. Pathological central fatigue is found in Multiple Sclerosis, Traumatic Brain Injury, Parkinson's Disease, and many others.

Mechanisms of peripheral fatigue are usually attributable to a neuronal or muscular origin. Peripheral fatigue results from a lack of response in the neuromuscular system after central stimulation [27]. Peripheral fatigue is characterized by the failure to sustain the force of muscle contraction [9]. Pathological peripheral fatigue is found in neuromuscular disorder, rhabdomyolysis, muscle ischemia, restless legs and more.

In many of the previously mentioned health conditions, physical inactivity is a contributing factor to the increased fatigue of the patient [30]. Deconditioning, as a result of restricted physical activity, results in large decreases in muscle mass and strength, as well as increased fatigue due to changes in muscle metabolism [31], [32]. Physical fatigue is also increasingly observed as a secondary outcome in many diseases and health conditions during the performance of everyday activities [32].

Non-Neurological Fatigue

The exact mechanism of how non-neurological disease causes fatigue is not fully understood [7]. However, there are indications that peripheral proinflammatory cytokines signal the central nervous system to initiate fatigue [33]. A common non-neurological cause of temporary fatigue is an infection or the common cold. Non-neurological causes of chronic fatigue include infectious diseases (human immunodeficiency virus, mononucleosis, Borreliosis, and chronic pancreatitis), hematologic disease (anaemia and hemochromatosis), dehydration, immunological disease (celiac disease), rheumatological disease, cardiac disease (heart failure and cardiomyopathy), endocrinologic disorder (diabetes, Addison's disease, hypopituitarism, and hypothyroidism), renal disease (insufficiency and dialysis), lung disease (chronic obstructive lung disease and asthma), malnutrition (poor diet, irritable bowel disease, eating disorders and hypoproteinemia), liver disease, chronic pain, chronic fatigue syndrome, fibromyalgia, malignancy (cancer, sarcoma, lymphoma, and leukaemia), Gulf War disease, poisoning, mineral or vitamin deficiencies, drugs, or irradiation [7].

Drugs and medications may also be a cause of non-neurological fatigue. The drugs that cause fatigue include alcohol, antihistamines, benzodiazepines, antispasmodics, antiepileptic drugs, neuroleptics, and narcotics [7].

Mental Fatigue

Mental fatigue in the pathological domain includes cognitive and affective (psychological/ psychiatric) fatigue. Cognitive fatigue has been studied in the context of MS [34], cancer [35], TBI [36], HIV [37], and other diseases. Affective fatigue is influenced by psychological factors (attitude, motivation, will, endurance, flexibility, inertia, persistence, concentration, and alertness) as well as psychiatric factors (depression, mania, psychosis, and addiction) [28]. Individuals with chronic fatigue report poorer mental health than their non-chronic fatigue counterparts [38].

Non-pathological Fatigue

In contrast to pathological fatigue, non-pathological fatigue is short term and remits with rest. Non-pathological fatigue is sometimes referred to as physiological fatigue in the scientific literature. Non-pathological fatigue alerts the individual to opportunity costs of current activities, and of the attraction of neglected needs and alternative goals [39]. Fatigue in healthy individuals is a universal experience and a natural occurrence after physical or mental efforts, usually relieved by rest. Research has examined biological explanations for pathological versus non-pathological fatigue [40], as well as self-report scales to distinguish fatigue associated disease from fatigue associated with healthy controls [41]. It has been reported that 55% of healthy individuals identified a physical sensation of fatigue and 24% identified a mental sensation of fatigue [26].

Physical Fatigue

From a physical perspective, fatigue is described as the inability of the muscles to maintain the required level of strength during exercise activities [42], [43]. It can also be characterized as an exercise-induced reduction in muscle's capability to generate force. There is no single cause of physical fatigue [44] and physical fatigue includes both central and peripheral fatigue.

Central fatigue designates a decrease in voluntary activation of the muscle, whereas, peripheral fatigue indicates a decrease in the contractile strength of the muscle fibres and changes in the mechanisms underlying the transmission of muscle action potentials [45]. Central and peripheral fatigue is a common experience during sport and exercise activities.

The impact of physical fatigue on cognitive performance depends both on the intensity and the duration of the exercise [46], [47]. Prolonged physical exercise leading to dehydration and physical fatigue is associated with a reduction in cognitive performance [48].

Mental Fatigue

Mental fatigue includes cognitive and affective fatigue and is an unfocused mental state, characterized by distraction, frustration, or discomfort. Mental fatigue is a psychobiological state caused by prolonged periods of demanding cognitive activity and characterized by subjective feelings of "tiredness" and "lack energy" [4].

In terms of cognitive activities, mental fatigue may be defined as the perception of feeling cognitively fatigued after performing demanding cognitive activities that involve concentration, attention, endurance, or alertness [49]. In the cognitive domain, fatigability can be measured as a decline in the reaction time, a decline in accuracy on continuous performance tasks, or a probe task that is given before and immediately after a fatiguing cognitive task [50], [51]. This cognitive fatigue is associated with problems completing tests, particularly where there is a requirement to sustain high levels of effort over time [39]. The effects of mental fatigue on cognitive performance [4], [51]–[53], and the skilled performance of drivers [54] and air pilots [55], have been investigated. Mental fatigue also limits physical performance [56] through perceived exertion [5]. Similarly, mental fatigue, following the performance of cognitive tasks, impairs emotion regulation [57].

Affective fatigue is characterized by low mood, tiredness, weariness, and lethargy [39]. It has been reported that 21% of healthy individuals identified an affective sensation of fatigue [26]. Non-pathological affective fatigue includes self-regulatory fatigue, empathy fatigue, and other fatigue associated with emotional depletion (burnout).

Factors Influencing Fatigue

Pathological and non-pathological fatigue is influenced by numerous factors, such as age, gender, physical condition, diet, latency to last meal, mental status, psychological conditions, personality type, life experience, and the health status of the individual [7]. Most studies found

more fatigue in women than in men [38], [58]–[62]. Inconsistent findings have been reported regarding age and fatigue [38], [58], [62], [63]. Additionally, a high level of formal education has been associated with a lower prevalence of fatigue [61], [64], [65].

Sleepiness and fatigue are distinct and interrelated. Sleepiness refers to an increased propensity to fall asleep [66], while fatigue refers to tiredness resulting from exertion or illness. Fatigue may be regarded as a motivational drive to rest [67] and non-pathological fatigue will usually remit with rest. Sleepiness is related to circadian and homeostatic influences and remits after sleep [68], but not after rest.

Energy and Fatigue Measures

Fatigue perception is frequently measured by self-report scales, while fatigability is frequently assessed by performance, capacity, and technology-reported measures [143]. Subjective measures include scales and prompts for assessment while objective measures include performance and capacity tasks (physical and cognitive), physiological measurements (cardiac, ocular, neural), and markers (biological and behavioural).

Subjective Measures

Fatigue perception is frequently measured by application of patient-reported outcomes (PRO) [12] through validated scales prompted for assessment. These scales may be administered momentarily, daily, monthly etc. through paper, web, or smartphone.

Scale Instruments

Scales for self-reporting may be unidimensional, evaluating a single property, or multidimensional, evaluating multiple properties [49]. These instruments address different aspects of fatigue and energy and some address more than one aspect. No single measure of fatigue adequately captures the complexity of the phenomenon [15]. Researchers have pointed out that “in developing fatigue scales, there is a ‘‘catch 22’’ situation: before a concept can be measured, it must be defined, and before a definition can be agreed upon, there must exist an instrument for assessing phenomenology. There is, unfortunately, no ‘‘gold standard’’ for fatigue, nor is there ever likely to be” [13]. Table 2 in this section depicts several scale instruments routinely used to measure energy and fatigue. The majority of these energy/fatigue self-report scales were designed for pathologic populations, but have been applied to non-pathologic populations as well.

Table 2. Scale instruments routinely used to measure energy and fatigue.

Instrument	**Recall period**	**Measures**	**Administration**	**Usage**
Fatigue	Usually	Unidimensional;	10 items, 5-level	Pathologic and

Assessment Scale (FAS) [69]	(“refer to how you usually feel”)	fatigue severity	Likert scale	non-pathological (developed for chronic fatigue)
Functional Assessment of Chronic Illness Therapy (FACIT-F) Fatigue Subscale [70]	Past week	Unidimensional; general fatigue	13 items, 5-level Likert scale	Pathologic (people with various chronic illnesses, including cancer)
Fatigue Impact Scale (FIS) [71]	Past month, present time	Multidimensional; physical, cognitive, and psychosocial functioning, total fatigue	40 items, 5-level Likert scale	Pathologic (developed for infectious disease patients)
Fatigue Questionnaire / Fatigue Scale (FQ / FS) [72]	Past month	Multidimensional; physical, mental, total, substantial, transient, and chronic fatigue	11 items, 4-level Likert scale	Pathologic and non-pathological (developed for use in hospital and community populations)
The Fatigue Severity Scale (FSS) [73]	Past week	Unidimensional; fatigue severity	9 items, 7-level Likert scale	Pathologic and non-pathological (developed for patients with Multiple Sclerosis or systemic lupus erythematosus)
Multidimensional Assessment of Fatigue (MAF) [74]	Past week	Multidimensional; degree, severity, distress, and impact of fatigue	16 items, 4-10-level Likert scales	Pathologic and non-pathological (developed for patients with rheumatoid arthritis)
The Multidimensional Fatigue Inventory (MFI) [75]	Lately (“refer to how you have been feeling lately”)	Multidimensional; physical, mental, and general fatigue; reduced activity and motivation	20 items, 7-level Likert scale	Pathologic and non-pathological (used in chronically unwell and well populations)
Medical Outcomes	Past month	Multidimensional;	4 items, 3-6-level	Pathologic and

Study Short Form (SF-36) Energy and Fatigue subscale [76]		physical, cognitive, social, and emotional functioning	Likert scales and yes/no	non-pathological (developed to measure the health status of individuals living in the community)
Patient-Reported Outcomes Measurement Information System (PROMIS), Fatigue short form or computerized adaptive test [77]	Past week	Multidimensional; physical, mental, general, emotional, total, substantial, transient, chronic fatigue; reduced activity and motivation; physical, cognitive, psychosocial, social, emotional functioning; energy	Up to 95 items, 5-level Likert scale	Pathologic and non-pathological (can reliably estimate fatigue reported by the U.S. general population)
Profile of Mood States (POMS), Fatigue and Vigour subscales [78]	Past week, present time	Multidimensional; physical and mental fatigue; energy	65 items, 5-level Likert scale	Non-pathological (adult version and adolescent version)
Visual Analog Scale to Evaluate Fatigue Severity (VAS-F) [79]	Present time: “right now”	Bidimensional; energy and fatigue	18 items, visual analogue	Pathologic and non-pathological (validated with adults aged 18–55 years)

Considerations in choosing a particular scale include recall period, unidimensionality or multidimensionality, scale structure and length, and suitable population. Scales differ in their scope, some measuring severity only, and others duration and impact on a range of functions [14]. Fatigue measures have been evaluated for the number of symptoms assessed, dimensions of fatigue explored, the time frame of the assessment, scale, method, the population on which the scale was developed, and psychometric properties [13], [14].

Some applications of these scales are illustrated below. SF-36 and PROMIS have been used in traditional studies assessing fatigue in the general population [14], [80]. FQ, FSS, and MAF have been employed to assess workplace-related fatigue [81], [82]. POMS has been used to assess fatigue in bus drivers [83] and sport athletes [24]. Scales were also used in traditional studies to assess energy and fatigue in individuals with a plethora of diseases, e.g., cancer [84], [85], cardiovascular disease [86], [87], chronic obstructive pulmonary disease [88], diabetes [89], fibromyalgia [90], hearing loss [16], inflammatory bowel disease [91]–[93], lupus [94], major

depressive disorder [95], multiple sclerosis [96]–[98], psoriasis [99], pulmonary arterial hypertension [100], renal disease [101], rheumatic disease [102], [103], sleep apnea [104], stroke [105], [106], and traumatic brain injury [10].

Smartphone collection of self-reported energy and fatigue data has been utilized in the context of multiple sclerosis [107], cancer-related fatigue [108], and bipolar disorder [109]. Smartphone data collection often incorporates validated scales. For example, a mobile phone application to collect data on self-reported fatigue for multiple sclerosis [107] incorporated PROMIS. Researchers concluded that a phone application incorporating PROMIS may be useful to provide estimates of fatigue to facilitate clinical monitoring of fatigue for clinic settings.

Momentary Assessments

The Ecological Momentary Assessment (EMA) is a technique that elicits a repeated, real-time measurement of behaviours or experiences as they occur in the naturalistic setting of an individual's daily life. This method was originally developed to perform *in situ* data collection for behavioural medicine [110]. The Experience Sampling Method (ESM) aims to assess participant thoughts, behaviours, and feelings during daily life by collecting self-reports, triggered at various moments during the day [111]. The two terms (EMA and ESM) are used interchangeably, and in practice, they are measured using the same methods [112].

Traditional studies employing EMA/ESM assessed fatigue and fatigability in segments of the general population. For instance, this method has been applied to demographic groups, work settings, and disease populations. Specifically, the relationship between women's passion for physical activity and vitality was examined using SF-36 scale [113]. Researchers have also employed POMS scale to examine occupational energy management strategies by hourly diary questions in academic workers [114]. A separate study examined the effects of breaks on regaining vitality in the workplace using an activation–deactivation adjective checklist [115]. Additionally, EMA/ESM assessment of energy/fatigue has been applied to disease populations including osteoarthritis ([116] researchers used SF-36 scale), kidney disease ([45] researchers used Daytime Insomnia Symptom Scale), and cancer ([117] researchers used a single-item fatigue intensity scale; [20] researchers used 10-point Likert scale for current fatigue).

Mobile-administered EMA/ESM has been applied to the management of diseases. For cancer and its treatment, fatigue is one of the most common and distressing side effects. Cancer-related fatigue causes disruption in all aspects of Quality of Life and may be a risk factor for reduced survival [118]. A mobile phone-based, symptom management system can assist in the management of chemotherapy-related toxicity in patients with breast, lung and colorectal cancer [108]. This system prompts patients to complete an electronic symptom questionnaire on their mobile phone twice a day. A systematic review of mobile apps for bipolar disorder [109] identified thirty-five symptom monitoring apps aiming at assisting users with symptom tracking.

Objective Fatigue and Energy Measures

Fatigability is primarily measured by quantifying the decline in one or more aspects of performance during the continuous performance of a prolonged task or comparing performance before and immediately after a prolonged performance of a separate fatigue-inducing task [8]. In pathological cases, individuals may experience fatigue even in mundane situations, such as daily activities [119]. When objectively measuring fatigue, it is important to indicate the domain examined and the task used to induce fatigability.

Fatigue-related decrements in task performance can be measured by following two common approaches. Ackerman [120] provides a classification of procedures for cognitive fatigue, however, we argue that these same approaches pertain to physical fatigue as well. The indirect approach consists in the assessment of cognitive ability before and after a prolonged period of time during which effort may vary. The direct approach consists in the continuous measurement of fatigue during the difficult task. The benefits of the first method are that all participants can complete the same task, while the variation lies in the difference between ex-ante and ex-post fatigue among individuals. This method does not quantify the performance decrease as a continuous function of time. Conversely, the second method can monitor fatigue accumulation, but the tasks may vary. One example is vigilance tasks, where participants are required to maintain attention for target events while ignoring other stimuli [121], [122].

There is a distinction between capacity (describing a person's ability to execute a task in a standardized, optimized, or controlled environment), capability (describing what a person can do in their daily environment), and performance (describing what a person actually does in their daily environment) [123]. Capacity is the composite of all the physical and mental capacities that an individual can draw on and performance is what individuals do in their current environment, including their involvement in life situations [124]. It is beyond the scope of this chapter to classify past studies as capturing capacity, capability, performance, or a combination.

Physical Assessment

The monitoring of fatigue and energy has been examined as an approach to maintain health, assist in disease management, and improve performance, productivity, and safety. A plethora of methods have been employed in order to monitor fatigue and energy: performance-reported outcomes (PerfRO) [12] for physical and cognitive fatigability, and tech-reported outcomes (TechRO) [12] from physiological processes (cardiac, ocular, neural) and markers (biologic, behavioural).

Fatigability is usually quantified as a decline in peak force (torque), power (velocity of muscle contraction), speed, fatigue index (force change over time), sense of effort, perception of effort, or accuracy of performance after performing a task, which requires physical effort [7]. Characteristics of tasks include exercise type, intensity, load, tested muscle, and physical environment [28].

The first dimension of physical performance fatigue is “physical capacity” (i.e., maximum performance). The two most common indicators of physical capacity are (1) the aerobic capacity and (2) the power output capacity. Measures of aerobic capacity include the maximal oxygen volume (VO_2-max). Measures of power output include the peak power output. Momentary exercises leading to the assessment of these measures include aerobic and resistance training [97]. Example exercises routinely used, e.g., in professional sports players include various jump protocols, including squat and countermovement jumps, which can lead to indirect assessments of fatigability [24]. Direct measures of fatigue include a joint range of motion or flexibility of appendages such as the knee, hip, groin, and other joints during the exercise.

The second measured dimension of fatigue is “muscular strength.” Studies measuring muscular strength included momentary resistance training of various types (weight machines, free weights, resistance bands, cycling ergometers) and other strength training (specialized locomotor training, cycling, aquatics) [97], muscular oxygen consumption (mVO_2), or electromyography (EMG).

The third dimension of fatigue is “mobility,” which is more commonly measured in cases of pathological fatigue. Mobility measures include the momentary 6-Minute Timed Walk (6MTW) [125], the Timed 25-Foot Walk [126], and the Timed Up & Go [127].

Exercise-specific hardware used for such exercises include treadmills, weight machines, free weights, and resistance bands. Technology-enhanced exercises include the robotic-assisted treadmill and functional electrical stimulation-assisted cycling [97], and transcranial magnetic stimulation [7]. Figures 2 and 3 respectively depict hardware and tasks used to measure physical performance.

Figure 2. Hardware for physical performance: ergometer (Monark), treadmill (LifeFitness).

Figure 3. Tasks to measure physical performance: audio reaction timer (American Educational Products), visual reaction screen (Cambridge Cognition).

Studies assessing non-pathological physical performance as a proxy for physical fatigue involved segments of the general population, e.g., physical fatigue in young adults using POMS, trail-making test on an iPad and mVO_2 [128], physical fatigue during a sit-to-stand physical test by using EMG and accelerometer (Samsung) in the lab [129], or PhysioLab, a physiological computing toolbox measuring multiple signals (ECG, EMG, and EDA) to study cardiorespiratory fitness in elderly populations [130], all momentary.

Other observational studies assessed physical fatigue in a pathological context with individuals with health conditions or diseases; assessments include the effects of caloric restriction on cardiorespiratory fitness and fatigue in older adults with obesity by using graded exercise tests measuring VO_2-max [131], the differences in motor fatigue between patients with stroke and patients with multiple sclerosis by using self-reported SF-36 and 6MWT [132], physical fatigue in lumbar disc herniation by using EMG [133].

Continuous monitoring studies assessed the effects of disease on fatigue, e.g., a rehabilitation program on aerobic fitness, cancer-related fatigue, and quality of life using subjective MFI and objective energy expenditure armbands (SenseWear) [134], or the fatigue monitoring system (FAMOS) which can monitor physiological parameters from multiple sclerosis patients and controls, all pathological.

Cognitive Assessment

In the cognitive domain, fatigue leads to the degradation of cognitive performance [121], as reflected by degradations in verbal, visual, short, and long-term memory, processing speed, primary and divided attention, verbal fluency, motor speed, reading speed, visual scanning, orientation, calculation, success rate, and other measures.

Cognitive assessments were measured by using numerous momentary measures, which collectively assess the above degradations. Table 3 reviews several task-based tests yielding cognitive performance-reported outcomes [12]. Figure 4 depicts a few tasks measuring cognitive performance.

Table 3. Tasks and measures of cognitive performance

Task	Measures	Administration	Usage
Mini-Mental State Examination (MMSE) [135]	Orientation, short-term memory registration, attention, calculation, recall, language, and task reproduction	16 complex items: qualitative and quantitative questions	Elders, potentially pathologic
Trail Making Test (TMT) [136]	Visual search, scanning, processing speed, mental flexibility, and executive functions.	Two items: the participant connects circles denoted by numbers and letters in ascending order	Non-pathological
Selective Reminding Test (SRT) [137]	Verbal memory	One item: the participant recalls as many as possible of 12 dictated unrelated words	Non-pathological
Spatial Recall Test (SPART) [138], [139]	Visuospatial learning, the susceptibility of such learning to proactive and retroactive interference, and the ability to recall visuospatial information following a period of delay	One item: the participant recalls as many as possible of 10 checkers on a 36-checkers square board	Pathologic, multiple sclerosis
Symbol Digits Modalities Test (SDMT) [140]	Presence of organic cerebral dysfunction leading to neurological impairment	One item: the participant has 90 seconds to pair specific numbers with given geometric figures.	Pathologic, cerebral dysfunction
Paced Auditory Serial Addition Test (PASAT) [141]	Rate of information processing after recovering from trauma	Multiple items: the participant hears a series of digits, one every 3 seconds, and reports the sum of the last two digits.	Pathologic, multiple sclerosis
Word List Generation (WLG) [142]	Neuropsychological measures of verbal fluency	One item: the participant generates words from a restricted category (e.g., starting with S or denoting animals) in 60 seconds.	Pathologic, dementia, multiple sclerosis
Rey Auditory Verbal Learning Test (RAVLT)	Recent memory, verbal learning, susceptibility to interference, and retention	Multiple items: 15 nouns read aloud each second for 5 consecutive trials followed	Non-pathological

[143]	of information after a certain period of time during which other activities are performed	by participant recall.	
Simple Reaction Time Task (SRTT) [144]	Relationships between the deceleration of heart rate observed to anticipate both aversive and non-aversive stimuli, and several aspects of the somatic-motor activity.	One item: A square is shown on screen at different intervals. The participant selects a button to react to seeing the square.	Non-pathological
Psychomotor Vigilance Task (PVT) [23]	Impact of loss of sleep sustained wakefulness, and/or time of day on neurobehavioral performance	Multiple items: ranging up to 10 minutes, similar to the SRTT.	Non-pathological
Brief Repeatable Battery of Neuropsychological Tests [145]	Selective short-term memory, spatial recall, symbol digit modalities, paced auditory serial addition, and word list generation; first used for multiple sclerosis	Multiple tests: selective reminding test (SRT), spatial recall test (SPART), symbol digits modalities test (SDMT), paced auditory serial addition test (PASAT), delayed recall of the SRT, delayed recall of the SPART, and word list generation (WLG).	Pathologic, multiple sclerosis

Cognitive performance studies included fatigue assessment in non-pathological segments of the general population, e.g., alertness, vitality, and sleepiness by using Psychomotor Vigilance Task (PVT) and other tasks in different lighting settings [146], occupational fatigue, e.g., in healthcare and medical staff by using the rate of error [147], or airline pilots on the flight deck by using PVT [148].

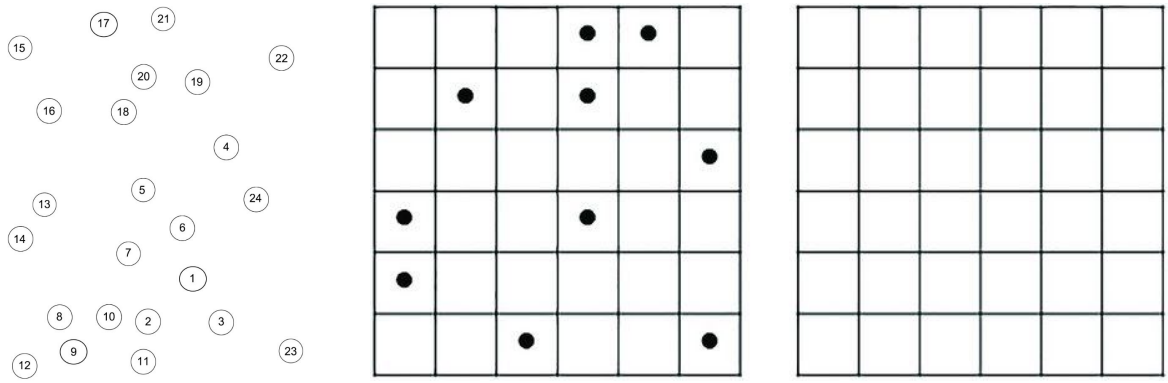


Figure 4. Tasks to measure cognitive performance: trail making test, spatial recall test.

Cognitive performance studies also included fatigue assessment in pathological settings, e.g., the relationships between health-related Quality of Life, fatigue, and exercise capacity in coronary artery disease individuals using MFI and a bicycle ergometer test [149].

Technology-driven studies include assessments of mental fatigue in a non-pathological context by performing tasks with a computer, e.g., keyboard and mouse interaction patterns [150] recovery from work exhaustion by use of Twitter [151], or in a pathological context. For example, those living with an acquired brain injury often have issues with cognitive fatigue due to factors resulting from the injury. Studies have shown fatigue to be one of the most disabling symptoms, regardless of the severity of brain injury [152]–[154]. Researchers presented a smartphone application for the evaluation of cognitive fatigue, which can be used daily to track cognitive performance in order to assess the influence of fatigue [155]. Researchers concluded that the presented smartphone application for the evaluation of cognitive fatigue could be utilized in everyday life.

Cardiac Physiology

Cardiac activity measures used to assess fatigue include the resting heart rate (HR), exercise heart rate (HRex), heart rate variability (HRV), and the heart rate recovery (HRR). The heart rate may increase or decrease in response to a variety of factors including physical and mental effort, distress, and anxiety that are potentially associated with fatigue [16]. Elevated HRV was observed during strenuous tasks in individuals with chronic fatigue [156] and healthy individuals of young age while performing a task [157]. HRR may serve as a marker of acute training-load alteration, however recent studies showed inconclusive results [24]. A more detailed measure of heart activity is the electrocardiogram (ECG), an electrophysiological method, which records the electric signals of the heart and from which the HR can be derived. Figure 5 depicts an electrocardiograph and electrode placement on the body.

Figure 5. Electrocardiography: electrocardiograph (Edan), electrode placement (Philips).

Studies using the ECG to assess fatigue in a non-pathological, occupational context include airline crew [158], surgeons [159] or 3D TV watchers [160]. In these studies, the ECG was measured with electrode-based devices before and after the tiring task (i.e., via an indirect measurement approach). HRV pre- and post-task was used as a measure for fatigue in work settings, e.g., emergency and pre-hospital doctors [161].

Measurements of cardiac physiology have been performed during daily life (i.e., via a direct measurement approach) also in a non-pathological setting. A large body of research focused on assessing cardiac activity in healthcare and driving professionals. Medical interns were given Holter recorders throughout the day, measuring HR and HRV, in conjunction with resting ECG to assess fatigue [162]. Surgeon HRV (using EEG) was assessed in robot-assisted versus conventional cholecystectomy [163]. Drivers were assessed while driving, through an ECG device mounted on the steering wheel [164]. Another study assessed the impact of electroacupuncture on fatigue and Quality of Life using subjective SF-36 and objective HRV using ECG (SphygmoCor) [165]. A method aimed at estimating the perception of physical fatigue by predicting heart rate through smartphones has been proposed by estimating the oxygen consumption, using a smartphone acceleration and location (via accelerometer and GPS, respectively) [3]. The study yielded an adequate detection of fatigue when individuals performed daily-life activities under naturalistic conditions.

Ocular Physiology

Keeping the eye closed or having fixed changes in pupil diameter have been observed in a state of fatigue [166] due to monotony or sleep deprivation. Ocular physiology measures used for assessing fatigue include the spontaneous eye blink [167], pupil diameter [168], oscillations in pupil diameter (fatigue waves) [169], [170]. Another method used to detect fatigue is the electrooculogram (EOG), an electrophysiological method, which measures the resting electrical potential between the cornea and Bruch's membrane.

Studies using ocular physiology measures were primarily done to assess fatigue in non-pathological, occupational settings, e.g., in the military detecting sleep deprivation-induced fatigue by saccade peak velocity in the Navy using questionnaires (on PDAs), actigraphy

(Actiwatch), and EOG (Natus, then Embla) during a saccade task [171] or assessing fatigue in the Air Force through saccadic velocity using software (Eyelink) in a dark room before and after a long flight [22]. For driver drowsiness, studies assessed fatigue by EOG using a device mounted next to the eyes for brief periods [21].

Smartphones have been utilized and applied to drivers as well. Researchers have presented an app, which uses information from both front and back cameras and others embedded sensors on the phone to detect and alert drivers to dangerous driving conditions inside and outside the car [172]. Researchers used computer vision and machine learning algorithms on the phone to monitor and detect whether the driver is tired or distracted using the front camera while at the same time tracking road conditions using the back camera. The front camera pipeline tracks the driver's head pose and direction as well as eyes and blinking rate as a means to infer drowsiness and distraction. Specifically, researchers used blink detection algorithms to detect periods of micro-sleep, fatigue and drowsiness. A more recent study improved EOG by mounting the device on the forehead to increase the duration of comfortable measurement [173].

Neural Physiology

Neural electrophysiological measures used to assess fatigue include the electroencephalogram (EEG), the evoked response potential (ERP), the Error Related Negativity, and lateralized readiness potential [16]. Magnetic resonance imaging (MRI) was also used to identify factors of fatigue [174]. This type of objective measure focuses on cognitive performance, described in a preceding section, by requiring the participants to conduct a task while monitoring takes place.

Studies have assessed neural physiology of fatigue in a non-pathological context by using EEG or ERP in the general population [175], [176], as well as EEG on occupational fatigue, e.g., drivers [177], [178], and surgeons while conducting a demanding task. For surgeons, Kahol [179] studied the impact of fatigue in surgical residents, which used a demanding task and measurement by EEG using a B-Alert device while Guru [180] assessed cognitive performance during robot-assisted surgery by EEG using a B-Alert device. Other studies which used electrophysiological measures in conjunction with other methods are elaborated on in the objective measures section of mix methods.

Biologic Markers

Fatigue-related biologic markers were studied in the pathological context of chronic disease: plasma glucose, associated with variations in transient physical and mental energy, effort, and fatigue with variable degrees of success [181], [182]; cortisol, an indirect marker of fatigue through stress level and energy expenditure associated with fatigue [183]; salivary alpha-amylase (sAA) associated with surrogate markers of nervous system activity [184] and task engagement/disengagement [185], with variable degrees of success; and melatonin following circadian patterns and disrupted in individuals with chronic disease and recurrent fatigue [186], used for sleep-related fatigue. In elite athletes, creatine kinase (CK), C-reactive

protein (CRP), uric acid, testosterone, salivary immunoglobulin (S-IgA) were used as indirect markers of fatigue in the recovery period following intense physical activity. Biologic systems involved in the regulation of motor activity are intricately linked with sleep, feeding behaviour, energy, and mood [187].

Behavioural Markers

Common behavioural markers utilized to assess fatigue include sleep and physical activity. These markers can be assessed by research-grade devices and consumer devices alike, with various degrees of validated accuracy, wear comfort, and presence in the research lab for the procedure. Figures 6 and 7 depict several research-grade and consumer wearable devices, respectively. As opposed to the momentary measures above, the behavioural markers can also be monitored continuously (with very high frequency, e.g., seconds or milliseconds) and longitudinally (for an extended duration, e.g., weeks to years) in time.

Sleep can be assessed using polysomnography and actigraphy. Polysomnography (PSG) [188] is an electrophysiological sleep study, which assesses brain waves (EEG), oxygen levels in the blood, heart rate (ECG), eye movements (EOG), and muscle and skeletal muscle activation and movements (EMG), breathing functions, respiratory airflow, respiratory effort, and pulse oximetry (SpO_2). Polysomnography quantifies sleep duration, interruptions, stages (e.g., light, deep, rapid eye movement (REM)) and waking states (e.g., awake, asleep). Actigraphy [189] is a non-invasive electrophysiological method that assesses movement and is used to monitor humans at rest or during various types of physical activity. Examples of research-grade wearable actigraph devices are ActiWatch[1] and ActiGraph[2]. The actigraph can be worn on the wrist or ankle during daily life, for several weeks. The actigraph allows for the continuous collection of data due to its non-invasive nature, however, widespread and longitudinal use is limited by its specific purpose of researching physical activity with limited considerations to the user experience and price.

[1] https://www.usa.philips.com/healthcare/product/HC1046964/actiwatch-spectrum-activity-monitor
[2] https://www.actigraphcorp.com

Figure 6. Research-grade wearables: Accusplit Pedometer AX2720MV (Accusplit), ActiGraph GT9X Link (ActiGraph), Actiwatch Spectrum Plus (Philips), ActTrust 2 (Condor Instruments), Embletta MPR Sleep System (Natus), Sensewear Bodymedia Fit (Sensewear).

More recent consumer wearable monitors, in the form of wristbands, smartwatches, sleep mattresses, or finger rings from manufacturers such as Fitbit[3], Oura[4], and Withings[5] [190] monitor sleep continuously by using a combination of movement, measured by a triaxial accelerometer, and HR/HRV, measured by photoplethysmography (PPG), non-invasive optical measurement of the volumetric variability of blood in the vessels under the skin. Consumer wearables can also measure behavioural markers pertaining to physical activity, e.g., duration, intensity (classified as, e.g., sedentary, low, moderate, and vigorous), type (using activity class recognition), effort (in metabolic equivalent of tasks (METs)), distance, elevation, step count, workouts, and other measures derived from the continuous multivariate data obtained from triaxial accelerometer and gyroscope sensors inside the device.

[3] https://fitbit.com
[4] https://ouraring.com
[5] https://withings.com

Figure 7. Consumer wearables: Apple Watch Series 5 (Apple), Fitbit Versa 2 (Fitbit), Garmin Fēnix 6 (Garmin), Huawei Watch GT 2 (Huawei), Oura Ring (Oura), Polar Ignite (Polar), Samsung Galaxy Active 2 (Samsung), Withings Steel HR (Withings), Fitbit Charge 4 (Fitbit), Garmin Vivosmart 4 (Garmin), Samsung Galaxy Fit (Samsung), Withings Pulse HR (Withings).

Studies assessing non-pathologic fatigue, sleep, and physical activity have been performed in segments of the general population and for several occupations, usually by combining subjective and objective measurements. In segments of the general population, Ellingson [17] studied the influence of active and sedentary behaviours on perceived energy and fatigue in women by using subjective POMS and SF-36 and objective physical activity by an accelerometer (Actigraph). For occupations, Rizzo [191] assessed the role of fatigue and sleepiness in drivers with obstructive sleep apnea by using subjective SF-36 and objective PSG. De Araújo Fernandes Jr. [192] quantified the impact of shift work on train drivers by using PVT and actigraphy (Actiwatch). Fernandes-Junior [192] assessed sleep, fatigue, and Quality of Life in night shift workers using subjective scale and actigraphy (Actiwatch). Towards the pathologic type of fatigue, Campbell [193] assessed fatigue and sleep in individuals having unexplained chronic fatigue by using subjective scales and objective PSG; Maher [194] quantified the relationships between fatigue, physical activity, and socio-demographic characteristics in children and adolescents with physical disabilities by using objective physical activity measurement using an accelerometer (Actigraph).

Numerous other studies have assessed pathologic fatigue in the context of a specific disease using PSG or actigraphy. Attarian [195], Kaynak [196], Veauthier [197], and Kaminska [198] studied relationships between sleep and fatigue in multiple sclerosis patients. Keefer (2006) and Shitrit [199] assessed sleep and fatigue in inflammatory bowel disease. Merikangas [187] used

a combination of EMA and actigraphy to assess energy, mood, and activity in individuals with depressive disorders. Sun [200] assessed the relationships between daytime napping and fatigue and Quality of Life in cancer individuals by using subjective scale and objective sleep quality (Actigraph). Ancoli-Israel [201] assessed sleep, fatigue, and circadian activity in women with breast cancer by using subjective scale and objective circadian rhythms using actigraphy (Actiwatch). Holliday [202] assessed fatigue and sleep quality in prostate cancer patients by using a subjective scale of Quality of Life and actigraphy (Actiwatch). Cambras [203] studied circadian rhythm in patients of encephalomyelitis using actigraphy (ActTrust). Nicklas [204] assessed physical activity behaviours (using accelerometers) and fatigue (using SF-36) in adults of middle and old age with chronic inflammations. Nilsson [205] studied intensity levels of physical activity and fatigue in cancer patients by using an accelerometer (SenseWear). Vancampfort [206] studied the relationships between cardio-respiratory fitness and increased quality of life in people with bipolar disorder using, among others, the subjective SF-36 and an armband (SenseWear) for objective physical activity and sedentary behaviour measurement. Sheshadri [207] assessed the relationship between intensity levels of physical activity and fatigue in patients on dialysis by using step count from a pedometer (Accusplit).

More recent studies used wearable to assess wearable-measured sleep and physical activity in a pathologic context. Qazi [208] studied fatigue in patients with inflammatory bowel disease by using a Fitbit Charge HR. Sofia et al. [209] used the same wearable to associate sleep fragmentation with individuals having clinically active disease. Abbott [210] conducted an intervention study for physical activity in case of cancer-related fatigue patients by using activity trackers (undisclosed brand) without reporting measurements but reporting that the activity tracker was deemed helpful.

Mixed Methods

In our literature review, we identified numerous studies which combined two or more objective measures of fatigue. These studies focused on either cognitive or physical fatigue in the general population or specific occupations, or physical fatigue in specific segments of the population.

For non-pathologic cognitive fatigue in the general population, Zhang [173] estimated mental fatigue based on EEG (Neuroscan) and HRV from ECG while performing an arithmetic task using a personal computer, Ren [211] studied various degrees of mental fatigue by using multiple types of measurements: EEG, ECG as well as galvanic skin response (GSR), Smith [212] quantified the effects on cognitive tasks on mental fatigue indicators, using PVT and other two tasks and assessing fatigue through subjective VAS and objective HRV from EEG, and Brown [213] studied the effects of mental fatigue on exercise intentions and behaviour using cognitive and then physical exercises by using a cycle ergometer.

In the area of non-pathologic physical fatigue, Kanitz [214] assessed the impact on eurythmy therapy on fatigue by using subjective MFI and objective HRV by ECG. For occupational fatigue, Smolders [215] studied the alertness during office hours induced by higher luminosity by using subjective measures, task performance (PVT, letter substitution test), and heart rate

measures (ECG), Oriyama [216] studied fatigue in shift nurses by measuring objective HRV from ECG, and subjective EMA using VAS, and Singh [217] assessed the technical performance of surgeons when using robotic surgery where the task was a suture under time pressure, measured with a subjective surgical task scale and objective HR, and objective functional near-infrared spectroscopy (fNIRS).

In the area of pathological fatigue, Dishman [218] studied the effects of cycling exercise on fatigue among young adults who report persistent fatigue using incremental exercise test on an electronically braked, computer-driven cycle ergometer (Lode), and providing subjective POMS and objective HR (Polar), VO_2-max and expired gas (Parvo Medics), and EEG (Electrical Geodesics).

Property Spectrums of Energy and Fatigue Measures

The findings from our literature review classify the energy and fatigue measurements by type (subjective and objective), location (clinician's office, daily life, or both/mixed), source (self-, performance/capacity-, and technology-reported, using the taxonomy by Mayo [12]), and administration (scales, prompts, tasks, and devices). We place each such measurement on spectrums for the following properties:

1. **Validated**: fatigue outcome reliability assessed by statistical analysis on the target population and scientific publication.
2. **Quantifiable**: fatigue outcomes interval or ratio at a minute or higher precision.
3. **Frequent**: often repeated administrations with one day or less between administrations.
4. **Continuous**: fatigue proxy variable measured on a time series with a minute or higher granularity.
5. **Judgment-free**: bias-free from the perception of judgment from the administrator; tasks and research devices allow some refraining.
6. **Mood-free**: bias-free from the voluntary or involuntary perception of self.
7. **Memory-free**: bias-free from the remembrance of the past; prompts allow for long-term memory loss.
8. **Owned**: whether the participant owns the device; scales and prompts are marked as partial in case they are delivered to a device owned by the participant.
9. **Contextual**: collected from settings daily life; research devices can be borrowed to the participant for a short time to wear in daily life context.

Table 4. Fatigue measurements and spectrums of characteristics from the literature review

Measurement	**Subjective**		**Objective**		
Location	Both office and daily life	Daily life	Office	Both office and daily life	Daily life

Reporting	Self-reported		Perf-reported	Tech-reported	
Administration	Scales	Prompts, e.g., EMA	Task hardware and devices	Research devices	Consumer devices
Validated	Yes	Partial	Yes	Yes	Partial
Quantifiable	No	No	Yes	Yes	Yes
Frequent	No	Yes	No	No	Yes
Continuous	No	No	Yes	Yes	Yes
Judgment-free	No	No	Partial	Yes	Yes
Mood-free	No	No	Yes	Yes	Yes
Memory-free	No	Partial	Yes	Yes	Yes
Owned	Partial	Partial	No	No	Yes
Contextual	No	Yes	No	Partial	Yes

Discussion

Key Findings

Fatigue or lack of energy is a universal symptom experienced by those suffering from different medical and psychological illnesses as well as by healthy individuals in the general population. Overall, fatigue is a ubiquitous and multifaceted symptom that is challenging to define and measure. Fatigue may be classified as pathological or non-pathological, physical or mental, and can be measured subjectively or objectively.

Different approaches have been employed in order to measure energy and fatigue including scales, prompts, physical measures, cognitive measures, physiological markers, biological markers, behavioural markers, and mixed methods. Some measurement methods assess the effects of fatigue (e.g. performance decrements), some attempt to identify the source of fatigue (e.g. muscle dysfunction), while others adopt a behavioural perspective (e.g. decreased physical activity or prolonged sleep). Some methods focus on capacity while others assess performance. These varied methods each contain advantages and disadvantages in terms of traditional validation, access to continuous data, and ecological validity.

Subjective instruments instantiating self-reported outcomes [12] suffer from inherent shortcomings, in particular, they are infrequent and subjective. Furthermore, self-report by recall has an intrinsic problem: due to biases, such as mood states or sleepiness, individuals are not able to accurately recall past experience, particularly experiences that are frequent, mundane, and irregular [219]. In addition, the potential discrepancy between how one feels and how one thinks one should feel contributes to lack of ecological validity in self-reports of fatigue and requires further research [15]. Incorporating a real-time collection of fatigue data in naturalistic settings may reduce problems associated with retrospective recall of events, summarization of events, and artificial contexts or settings [117].

Objective measures obtained by tech-reported outcomes can be collected continuously from individuals in the context of daily life. To this end, both academia and industry are increasing their efforts to develop technological solutions, such as sensors which can measure, models which can assess, and artefacts which can manage energy and fatigue. Recent technological methods to monitor and manage energy and fatigue include sensors, smartphones and their applications, and research- and consumer-grade wearables. Technology-based monitoring of energy and fatigue could assist in the initial diagnosis and the early detection of diseases could enable one to monitor post-treatment evolution and could help assess the risk of certain medications on patients [3]. Furthermore, technology-based monitoring of energy and fatigue could assist healthy individuals in enhancing work performance, conserving and managing energy levels, and maintaining health.

Energy and fatigue are of great importance to diseased individuals. The connection between pathological fatigue and disease is well established in the literature. Fatigue frequently foreshadows conditions like multiple sclerosis [220], cancer [28], and HIV infection [221], among other diseases. Furthermore, fatigue, as well as increased energy, has been identified as a core symptom of mental health disorders including depressive disorders and bipolar disorders. Current literature on energy and fatigue is biased towards pathological, rather than healthy, populations.

In addition to the comprehensive literature examining fatigue and disease, the monitoring of energy and fatigue has also been highlighted for specific vocational and occupational populations, such as professional athletes [24], police [25], and drivers [164]. The literature aims to gain an understanding of health, safety, occupational functioning, burnout, performance, and capacity. More efforts could be put toward studying healthy general populations, as in addition to affecting an individual’s quality of life, fatigue impacts the economy because of the connection to productivity and illness.

Insights into the classification and measurement of energy and fatigue may also be applied broadly to the general population as mobile monitoring technology allows the assessment of these homeostatic systems in real-time [187]. Quality of Life Technologies (QoLT) refers to technologies for assessment or improvement of the individual's quality of life [222]. Optimal measurement of energy and fatigue would be moved out of the lab and into the real world, being continuous rather than infrequent, and based on accurate, validated, yet minimally intrusive

measures and devices. Future research could establish traditional validity for the continuous, daily life, measurement of energy and fatigue.

Assessing energy and fatigue could also contribute to the quantified self. The quantified self (QS) is any individual engaged in the self-tracking of any kind of biological, physical, behavioural, or environmental information. QS promotes a proactive stance toward obtaining information and acting on it [223]. One of the earliest recorded examples of quantified self-tracking is that of Sanctorius of Padua, who studied energy expenditure by tracking his food intake, weight, and elimination for 30 years in the 16th century [224]. State of the art energy and fatigue assessment could contribute meaningfully to the quantified self.

Limitations

A limitation of the current chapter stems from the pathological bias in the field. Namely, because the existing literature is biased toward pathological fatigue, we built the non-pathological (also referred to as physiological) classification system arm based on existing pathological models. This limitation is also related to our literature search strategy. Our method of reviewing the literature was based on a scoping review approach rather than a structured systematic review. We did not exclude studies based on methodologies used or populations studied.

Subjective measures discussed in this chapter contain limitations including being infrequent, involving recalls, and potential to be influenced by mood states, memory, and expectations. Wearable measurements also contain limitations related to the population that uses wearables. Specifically, device owners are more likely to be young individuals with disposable incomes who already lead healthy lifestyles and want to quantify their progress [225]. Future work should ensure that wearable data is representative and note this bias in current wearable data.

An additional limitation of the field is that there is not yet a validated calibration between objective measures and the concept of energy and fatigue. Therefore, much of our discussion is speculative. A major impediment in the understanding of fatigue and energy lies in the fact that for over 100 years, research has shown little relationship between self-report and actual, objective measurements of fatigue [166]. There are several definitions of energy and fatigue and these have not been conclusively associated with objective measures. This doesn't invalidate subjective or objective measures of fatigue but rather indicates that they may be describing something that is more complicated and cannot be whittled down to a single biological measure. Therefore, both subjective experience and objective measurements are being considered in the context of energy and fatigue, as they are important indicators for health and quality of life. Future research could aim to bridge the gap between subjective and objective measures by accounting for multiple variables and conducting calibration studies.

Opportunities

Energy and fatigue is a Quality of Life facet in which the successful assessment, exclusively through Quality of Life Technologies [222], has promising likelihood. The mass adoption of miniaturized devices in daily life (with large scale and diversity in personal and contextual characteristics of the data), the availability of relevant predictors of energy and fatigue in large scale data, and the presence of platforms that facilitate participation in research at scale contribute to the feasibility of the operationalization of this facet.

Currently, research is progressing in assessing pathological and non-pathological energy and fatigue by using subjective, objective, and mixed methods. Miniaturized devices, such as smartphones and wearables, increasingly accurately monitor daily life behaviours (e.g., physical activity and sleep), sense signals (e.g., heart rate, momentary electrocardiogram, etc.) and administer prompts (e.g., validated scales, items, and tasks). As the line between consumer health wearables and medical devices continues to blur, it is possible for a single wearable device to monitor a range of medical risk factors [226]. Adoption of wearables is increasing; 21% of Americans own a wearable [227], there are more than 200 models of wearables[6] and the market is expected to continue to increase by 2022 [228] towards available objective behavioural data at scale. Open health platforms are being employed to facilitate scalable participation and manage subjective, objective, and mixed data [229].

Co-calibrations of (1) subjective validated scales of energy and fatigue and (2) objective measures of daily life behaviours may rigorously validate objective measures of energy and fatigue and meet the aim of assessing energy and fatigue using QoLT. For example, a study aiming to co-calibrate subjective scales and objective behaviours for occupational fatigue may collect multiple behavioural markers passively and continuously (e.g., physical activity, sleep, heart rate) from tens to hundreds of drivers for several months to years, during driving *and* daily living, and regularly administering validated energy and fatigue scales such that their recall periods cover the duration. Such a study may observe trends of fatigue longitudinally in time. Within a smaller sample size, a purely statistical approach would allow for the assessment of validity (e.g., by correlating the corresponding subjective and objective measures) and reliability (e.g., by measuring the same person's fatigue in similar days of week, months, or seasons) of the objective measure. Within a larger sample, a predictive approach would learn the subjective measures of energy and fatigue by using the objective measures of behaviours. These approaches can iteratively reduce the number of scale items. One step further, continuous behaviour monitoring during daily life facilitates the trigger of momentary assessments upon changes in objective behaviours that associate with changes in energy and fatigue. Such an approach may increase the accuracy of the co-calibration. Furthermore, alternative statistical or predictive risk scenarios can maintain energy (*"if you continue working at this pace, you will likely not get tired"*), prevent fatigue (*"if you continue working at this pace, you will likely accumulate occupational fatigue in two weeks"*), and compensate for the losses induced by

[6] https://www.inkin.com/wearables/

fatigue (*“you need to take a break of one week to restore your productivity from three months ago and compensate for the loss”*).

Initially, co-calibrations may suffer from lower accuracy (e.g., revealing only basic trends and associations) or limited extent (e.g., applying for specific scale items, collecting limited objective behaviours, applying for limited energy and fatigue types) as the measured objective measures or available sample may not explain the energy and fatigue directly. In such cases, a directed graph of co-calibrations with additional Quality of Life facets (e.g., stress, health outcomes), using additional objective measures, may need to be constructed to represent the relationships accurately such that energy and fatigue are explained through a series of directed co-calibration paths originating exclusively from objective measures, essentially assessing energy and fatigue through QoLT exclusively.

A successful energy and fatigue assessment using QoLT would contribute to the *“Internet of everything”* 50-year vision of a digital future where *“internet use will be nearly as pervasive and necessary as oxygen”* [230]. Specifically, such an assessment would contribute to three of Stansberry’s five hopeful visions of 2069. The first vision, *living longer and feeling better* where *“internet-enabled technology will help people live longer and healthier lives; scientific advances will continue to blur the line between human and machine”* [230] will be enabled by quantifying the relationships between energy, fatigue, behaviours, health, and Quality of Life outcomes. The second vision, *less work, more leisure* where *“artificial intelligence tools will take over repetitive, unsafe and physically taxing labour, leaving humans with more time for leisure”* [230] will be enabled through (short-term) the transition to increasingly passive reported outcomes that reduce the burden of participation in research and (longer-term) statistical and predictive optimization of physical and mental effort allocation for the occupations where energy and fatigue are prevalent. The third vision, *individualized experiences* where *“digital life will be tailored to each user”* [230] will be enabled by interventions leveraging large scale data, accurate models, and alternative personalized scenarios addressing fatigue prevention, before management, and before compensation.

Conclusive Remarks

Energy and fatigue impact physical, cognitive, emotional, social, and occupational functioning and carry important implications for an individual's health and overall Quality of Life. Lacking energy carries consequences for an individual’s routine functioning. Everyday activities, including work performance and self-care activities, can be impeded or even curtailed. Energy is required to sustain life and efficient spending of energy results in overall vitality.

The contributions of this chapter include a semi-structured literature review on energy and fatigue assessment and its potential within Quality of Life Technologies, a taxonomy of the field of energy and fatigue, and the identification of a research validation gap between subjective and objective measures of energy and fatigue. We foresee the necessity to conduct studies of

increasing size in order to co-calibrate the subjective and objective measures towards the integration of exclusively objective measures in research and clinical practice.

The measurement of energy and fatigue has been complicated by difficulties in definition and assessment. We conclude that optimal classification and measurement of energy and fatigue would occur in the real world, continuously and in real-time, while being ecologically valid and informing the design of interventions aimed at maintaining energy and monitoring fatigue towards positive outcomes of health and Quality of Life.

References

[1] "WHO | WHOQOL: Measuring Quality of Life," *WHO*. https://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/index4.html (accessed Jun. 24, 2020).

[2] S. M. Stahl, "The psychopharmacology of energy and fatigue," *J. Clin. Psychiatry*, vol. 63, no. 1, pp. 7–8, Jan. 2002, doi: 10.4088/jcp.v63n0102.

[3] N. Hernández and J. Favela, "Assessing the Perception of Physical Fatigue Using Mobile Sensing," in *Intelligent Data Sensing and Processing for Health and Well-Being Applications*, M. Wister, P. Pancardo, F. Acosta, and J. A. Hernández, Eds. Academic Press, 2018, pp. 161–173.

[4] M. A. S. Boksem and M. Tops, "Mental fatigue: Costs and benefits," *Brain Res. Rev.*, vol. 59, no. 1, pp. 125–139, Nov. 2008, doi: 10.1016/j.brainresrev.2008.07.001.

[5] S. M. Marcora, W. Staiano, and V. Manning, "Mental fatigue impairs physical performance in humans," *J. Appl. Physiol.*, vol. 106, no. 3, pp. 857–864, Mar. 2009, doi: 10.1152/japplphysiol.91324.2008.

[6] M. R. Al-Mulla, F. Sepulveda, and M. Colley, "A Review of Non-Invasive Techniques to Detect and Predict Localised Muscle Fatigue," *Sensors*, vol. 11, no. 4, Art. no. 4, Apr. 2011, doi: 10.3390/s110403545.

[7] J. Finsterer and S. Z. Mahjoub, "Fatigue in Healthy and Diseased Individuals," *Am. J. Hosp. Palliat. Med.*, vol. 31, no. 5, pp. 562–575, Aug. 2014, doi: 10.1177/1049909113494748.

[8] B. M. Kluger, L. B. Krupp, and R. M. Enoka, "Fatigue and fatigability in neurologic illnesses," *Neurology*, vol. 80, no. 4, pp. 409–416, Jan. 2013, doi: 10.1212/WNL.0b013e31827f07be.

[9] A. Chaudhuri and P. O. Behan, "Fatigue in neurological disorders," *The Lancet*, vol. 363, no. 9413, pp. 978–988, Mar. 2004, doi: 10.1016/S0140-6736(04)15794-2.

[10] T. Mollayeva, T. Kendzerska, S. Mollayeva, C. M. Shapiro, A. Colantonio, and J. D. Cassidy, "A systematic review of fatigue in patients with traumatic brain injury: The course, predictors and consequences," *Neurosci. Biobehav. Rev.*, vol. 47, pp. 684–716, Nov. 2014, doi: 10.1016/j.neubiorev.2014.10.024.

[11] E. Ream and A. Richardson, "Fatigue: a concept analysis," *Int. J. Nurs. Stud.*, vol. 33, no. 5, pp. 519–529, Oct. 1996, doi: 10.1016/0020-7489(96)00004-1.

[12] N. E. Mayo, S. Figueiredo, S. Ahmed, and S. J. Bartlett, "Montreal Accord on Patient-Reported Outcomes (PROs) use series – Paper 2: terminology proposed to measure what matters in health," *J. Clin. Epidemiol.*, vol. 89, pp. 119–124, Sep. 2017, doi: 10.1016/j.jclinepi.2017.04.013.

[13] A. J. Dittner, S. C. Wessely, and R. G. Brown, "The assessment of fatigue: A practical guide for clinicians and researchers," *J. Psychosom. Res.*, vol. 56, no. 2, pp. 157–170, Feb. 2004, doi: 10.1016/S0022-3999(03)00371-4.

[14] L. Whitehead, "The Measurement of Fatigue in Chronic Illness: A Systematic Review of

Unidimensional and Multidimensional Fatigue Measures," *J. Pain Symptom Manage.*, vol. 37, no. 1, pp. 107–128, Jan. 2009, doi: 10.1016/j.jpainsymman.2007.08.019.
[15] L. S. Aaronson *et al.*, "Defining and Measuring Fatigue," *Image J. Nurs. Sch.*, vol. 31, no. 1, pp. 45–50, 1999, doi: 10.1111/j.1547-5069.1999.tb00420.x.
[16] B. W. Y. Hornsby, G. Naylor, and F. H. Bess, "A Taxonomy of Fatigue Concepts and Their Relation to Hearing Loss," *Ear Hear.*, vol. 37, no. Suppl 1, pp. 136S-144S, 2016, doi: 10.1097/AUD.0000000000000289.
[17] L. Ellingson, K. Ae, N. Vack, and D. Cook, "Active and sedentary behaviors influence feelings of energy and fatigue in women.," *Med. Sci. Sports Exerc.*, vol. 46, no. 1, pp. 192–200, Jan. 2014, doi: 10.1249/mss.0b013e3182a036ab.
[18] K. Abdel-Kader *et al.*, "Ecological momentary assessment of fatigue, sleepiness, and exhaustion in ESKD," *BMC Nephrol.*, vol. 15, no. 1, p. 29, Feb. 2014, doi: 10.1186/1471-2369-15-29.
[19] A. Cochran, L. Belman-Wells, and M. McInnis, "Engagement Strategies for Self-Monitoring Symptoms of Bipolar Disorder With Mobile and Wearable Technology: Protocol for a Randomized Controlled Trial," *JMIR Res. Protoc.*, vol. 7, no. 5, p. e130, 2018, doi: 10.2196/resprot.9899.
[20] S. L. Curran, A. O. Beacham, and M. A. Andrykowski, "Ecological Momentary Assessment of Fatigue Following Breast Cancer Treatment," *J. Behav. Med.*, vol. 27, no. 5, pp. 425–444, Oct. 2004, doi: 10.1023/B:JOBM.0000047608.03692.0c.
[21] Thum Chia Chieh, Mohd. M. Mustafa, A. Hussain, S. F. Hendi, and B. Y. Majlis, "Development of vehicle driver drowsiness detection system using electrooculogram (EOG)," in *2005 1st International Conference on Computers, Communications, Signal Processing with Special Track on Biomedical Engineering*, Nov. 2005, pp. 165–168, doi: 10.1109/CCSP.2005.4977181.
[22] C. Diaz-Piedra, H. Rieiro, J. Suárez, F. Rios-Tejada, A. Catena, and L. L. D. Stasi, "Fatigue in the military: towards a fatigue detection test based on the saccadic velocity," *Physiol. Meas.*, vol. 37, no. 9, pp. N62–N75, Aug. 2016, doi: 10.1088/0967-3334/37/9/N62.
[23] D. F. Dinges and J. W. Powell, "Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations," *Behav. Res. Methods Instrum. Comput.*, vol. 17, no. 6, pp. 652–655, Nov. 1985, doi: 10.3758/BF03200977.
[24] R. T. Thorpe, G. Atkinson, B. Drust, and W. Gregson, "Monitoring Fatigue Status in Elite Team-Sport Athletes: Implications for Practice," *Int. J. Sports Physiol. Perform.*, vol. 12, no. s2, pp. S2-27-S2-34, Apr. 2017, doi: 10.1123/ijspp.2016-0434.
[25] Vila, Bryan, "Tired Cops: The Importance of Managing Police Fatigue," *Police Exec. Res. Forum*, 2000.
[26] A. Glaus, *Fatigue in Patients with Cancer: Analysis and Assessment*. Springer Science & Business Media, 2012.
[27] L. A. Jason, M. Evans, M. Brown, and N. Porter, "What is Fatigue? Pathological and Nonpathological Fatigue," *PM&R*, vol. 2, no. 5, pp. 327–331, May 2010, doi: 10.1016/j.pmrj.2010.03.028.
[28] M. P. Davis, D. Khoshknabi, and G. H. Yue, "Management of fatigue in cancer patients," *Curr. Pain Headache Rep.*, vol. 10, no. 4, pp. 260–269, 2006, doi: 10.1007/s11916-006-0030-2.
[29] M. Tanaka and Y. Watanabe, "Supraspinal regulation of physical fatigue," *Neurosci. Biobehav. Rev.*, vol. 36, no. 1, pp. 727–734, Jan. 2012, doi: 10.1016/j.neubiorev.2011.10.004.
[30] G. C. P. Bogdanis, "Effects of Physical Activity and Inactivity on Muscle Fatigue," *Front. Physiol.*, vol. 3, 2012, doi: 10.3389/fphys.2012.00142.
[31] Bloomfield, S. A., "Changes in musculoskeletal structure and function with prolonged bed rest," *Med. Sci. Sports Exerc.*, vol. 29, no. 2, pp. 197–206, 1997.
[32] J. H. Rimmer, W. Schiller, and M.-D. Chen, "Effects of Disability-Associated Low Energy Expenditure Deconditioning Syndrome," *Exerc. Sport Sci. Rev.*, vol. 40, no. 1, pp. 22–29, Jan. 2012,

doi: 10.1097/JES.0b013e31823b8b82.
[33] R. Dantzer *et al.*, “Identification and Treatment of Symptoms Associated with Inflammation in Medically Ill Patients,” *Psychoneuroendocrinology*, vol. 33, no. 1, pp. 18–29, Jan. 2008, doi: 10.1016/j.psyneuen.2007.10.008.
[34] S. Linnhoff, M. Fiene, H.-J. Heinze, and T. Zaehle, “Cognitive Fatigue in Multiple Sclerosis: An Objective Approach to Diagnosis and Treatment by Transcranial Electrical Stimulation,” *Brain Sci.*, vol. 9, no. 5, Art. no. 5, May 2019, doi: 10.3390/brainsci9050100.
[35] A. D. Valentine and C. A. Meyers, “Cognitive and mood disturbance as causes and symptoms of fatigue in cancer patients,” *Cancer*, vol. 92, no. S6, pp. 1694–1698, 2001, doi: 10.1002/1097-0142(20010915)92:6+<1694::AID-CNCR1499>3.0.CO;2-S.
[36] A. D. Kohl, G. R. Wylie, H. M. Genova, F. G. Hillary, and J. DeLuca, “The neural correlates of cognitive fatigue in traumatic brain injury using functional MRI,” *Brain Inj.*, vol. 23, no. 5, pp. 420–432, Jan. 2009, doi: 10.1080/02699050902788519.
[37] E. Byun, C. L. Gay, and K. A. Lee, “Sleep, Fatigue, and Problems with Cognitive Function in Adults Living with HIV,” *J. Assoc. Nurses AIDS Care JANAC*, vol. 27, no. 1, pp. 5–16, 2016, doi: 10.1016/j.jana.2015.10.002.
[38] W. S. Wong and R. Fielding, “Prevalence of chronic fatigue among Chinese adults in Hong Kong: A population-based study,” *J. Affect. Disord.*, vol. 127, no. 1, pp. 248–256, Dec. 2010, doi: 10.1016/j.jad.2010.04.029.
[39] B. Hockey and R. Hockey, *The Psychology of Fatigue: Work, Effort and Control*. Cambridge University Press, 2013.
[40] A. R. Light, A. T. White, R. W. Hughen, and K. C. Light, “Moderate exercise increases expression for sensory, adrenergic and immune genes in chronic fatigue syndrome patients, but not in normal subjects,” *J. Pain Off. J. Am. Pain Soc.*, vol. 10, no. 10, pp. 1099–1112, Oct. 2009, doi: 10.1016/j.jpain.2009.06.003.
[41] J. E. Schwartz, L. Jandorf, and L. B. Krupp, “The measurement of fatigue: A new instrument,” *J. Psychosom. Res.*, vol. 37, no. 7, pp. 753–762, Oct. 1993, doi: 10.1016/0022-3999(93)90104-N.
[42] R. H. Edwards, “Human muscle function and fatigue,” *Ciba Found Symp*, vol. 82, pp. 1–8, 1981.
[43] J. H. Friedman *et al.*, “Fatigue in Parkinson’s disease: A review,” *Mov. Disord.*, vol. 22, no. 3, pp. 297–308, 2007, doi: 10.1002/mds.21240.
[44] S. P. Cairns, A. J. Knicker, M. W. Thompson, and G. Sjøgaard, “Evaluation of Models Used to Study Neuromuscular Fatigue,” *Exerc. Sport Sci. Rev.*, vol. 33, no. 1, pp. 9–16, Jan. 2005.
[45] H. M. Abd-Elfattah, F. H. Abdelazeim, and S. Elshennawy, “Physical and cognitive consequences of fatigue: A review,” *J. Adv. Res.*, vol. 6, no. 3, pp. 351–358, May 2015, doi: 10.1016/j.jare.2015.01.011.
[46] K. Kamijo, Y. Nishihira, T. Higashiura, and K. Kuroiwa, “The interactive effect of exercise intensity and task difficulty on human cognitive processing,” *Int. J. Psychophysiol.*, vol. 65, no. 2, pp. 114–121, Aug. 2007, doi: 10.1016/j.ijpsycho.2007.04.001.
[47] P. D. Tomporowski, “Effects of acute bouts of exercise on cognition,” *Acta Psychol. (Amst.)*, vol. 112, no. 3, pp. 297–324, Mar. 2003, doi: 10.1016/S0001-6918(02)00134-8.
[48] C. Cian, P. A. Barraud, B. Melin, and C. Raphel, “Effects of fluid ingestion on cognitive function after heat stress or exercise-induced dehydration,” *Int. J. Psychophysiol.*, vol. 42, no. 3, pp. 243–251, Nov. 2001, doi: 10.1016/S0167-8760(01)00142-8.
[49] C. Falup-Pecurariu, “Fatigue assessment of Parkinson’s disease patient in clinic: specific versus holistic,” *J. Neural Transm.*, vol. 120, no. 4, pp. 577–581, Apr. 2013, doi: 10.1007/s00702-013-0969-1.
[50] J. Persson, K. M. Welsh, J. Jonides, and P. A. Reuter-Lorenz, “Cognitive fatigue of executive processes: Interaction between interference resolution tasks.,” *Neuropsychologia*, vol. 45, no. 7, pp. 1571–1579, Apr. 2007, doi: 10.1016/j.neuropsychologia.2006.12.007.

[51] D. van der Linden and P. Eling, “Mental fatigue disturbs local processing more than global processing,” *Psychol. Res.*, vol. 70, no. 5, pp. 395–402, Sep. 2006, doi: 10.1007/s00426-005-0228-7.
[52] M. M. Lorist, “Impact of top-down control during mental fatigue,” *Brain Res.*, vol. 1232, pp. 113–123, Sep. 2008, doi: 10.1016/j.brainres.2008.07.053.
[53] D. B. Cook, P. J. O’Connor, G. Lange, and J. Steffener, “Functional neuroimaging correlates of mental fatigue induced by cognition among chronic fatigue syndrome patients and controls,” *NeuroImage*, vol. 36, no. 1, pp. 108–122, May 2007, doi: 10.1016/j.neuroimage.2007.02.033.
[54] S. K. L. Lal and A. Craig, “A critical review of the psychophysiology of driver fatigue,” *Biol. Psychol.*, vol. 55, no. 3, pp. 173–194, Feb. 2001, doi: 10.1016/S0301-0511(00)00085-5.
[55] J. H. Goode, “Are pilots at risk of accidents due to fatigue?,” *J. Safety Res.*, vol. 34, no. 3, pp. 309–313, Aug. 2003, doi: 10.1016/S0022-4375(03)00033-1.
[56] J. Van Cutsem, S. Marcora, K. De Pauw, S. Bailey, R. Meeusen, and B. Roelands, “The Effects of Mental Fatigue on Physical Performance: A Systematic Review,” *Sports Med.*, vol. 47, no. 8, pp. 1569–1588, Aug. 2017, doi: 10.1007/s40279-016-0672-0.
[57] C. Grillon, D. Quispe-Escudero, A. Mathur, and M. Ernst, “Mental fatigue impairs emotion regulation,” *Emot. Wash. DC*, vol. 15, no. 3, pp. 383–389, Jun. 2015, doi: 10.1037/emo0000058.
[58] M. K. Chen, “The epidemiology of self-perceived fatigue among adults,” *Prev. Med.*, vol. 15, no. 1, pp. 74–81, Jan. 1986, doi: 10.1016/0091-7435(86)90037-X.
[59] G. E. Hardy, D. A. Shapiro, and C. S. Borrill, “Fatigue in the workforce of national health service trusts: Levels of symptomatology and links with minor psychiatric disorder, demographic, occupational and work role factors,” *J. Psychosom. Res.*, vol. 43, no. 1, pp. 83–92, Jul. 1997, doi: 10.1016/S0022-3999(97)00019-6.
[60] I. B. Hickie, A. W. Hooker, B. K. Bennett, D. Hadzi-Pavlovic, A. J. Wilson, and A. R. Lloyd, “Fatigue in selected primary care settings: sociodemographic and psychiatric correlates,” *Med. J. Aust.*, vol. 164, no. 10, pp. 585–588, 1996, doi: 10.5694/j.1326-5377.1996.tb122199.x.
[61] J. H. Loge, Ø. Ekeberg, and S. Kaasa, “Fatigue in the general norwegian population: Normative data and associations,” *J. Psychosom. Res.*, vol. 45, no. 1, pp. 53–65, Jul. 1998, doi: 10.1016/S0022-3999(97)00291-2.
[62] T. Pawlikowska, T. Chalder, S. R. Hirsch, P. Wallace, D. J. M. Wright, and S. C. Wessely, “Population based study of fatigue and psychological distress,” *BMJ*, vol. 308, no. 6931, pp. 763–766, Mar. 1994, doi: 10.1136/bmj.308.6931.763.
[63] A. David *et al.*, “Tired, weak, or in need of rest: fatigue among general practice attenders.,” *BMJ*, vol. 301, no. 6762, pp. 1199–1202, Nov. 1990.
[64] U. Bultmann, I. Kant, and S. V. Kasl, “Fatigue and psychological distress in the working population Psychometrics, prevalence, and correlates,” *U Bu*, p. 8, 2002.
[65] J. Van Mens-Verhulst and J. Bensing, “Distinguishing between chronic and nonchronic fatigue, the role of gender and age,” *Soc. Sci. Med.*, vol. 47, no. 5, pp. 621–634, Sep. 1998, doi: 10.1016/S0277-9536(98)00116-6.
[66] W. C. Dement and M. A. Carskadon, “Current Perspectives on Daytime Sleepiness: The Issues,” *Sleep*, vol. 5, no. suppl_2, pp. S56–S66, Sep. 1982, doi: 10.1093/sleep/5.S2.S56.
[67] A. Shahid, J. Shen, and C. M. Shapiro, “Measurements of sleepiness and fatigue,” *J. Psychosom. Res.*, vol. 69, no. 1, pp. 81–89, Jul. 2010, doi: 10.1016/j.jpsychores.2010.04.001.
[68] D. Dijk and C. Czeisler, “Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans,” *J. Neurosci.*, vol. 15, no. 5, pp. 3526–3538, May 1995, doi: 10.1523/JNEUROSCI.15-05-03526.1995.
[69] H. J. Michielsen, J. De Vries, and G. L. Van Heck, “Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale,” *J. Psychosom. Res.*, vol. 54, no. 4, pp. 345–352,

Apr. 2003, doi: 10.1016/S0022-3999(02)00392-6.
[70] K. Webster, L. Odom, A. Peterman, L. Lent, and D. Cella, “The Functional Assessment of Chronic Illness Therapy (FACIT) measurement system: Validation of version 4 of the core questionnaire,” *Qual. Life Res.*, 1999.
[71] J. D. Fisk, P. G. Ritvo, L. Ross, D. A. Haase, T. J. Marrie, and W. F. Schlech, “Measuring the Functional Impact of Fatigue: Initial Validation of the Fatigue Impact Scale,” *Clin. Infect. Dis.*, vol. 18, no. Supplement_1, pp. S79–S83, Jan. 1994, doi: 10.1093/clinids/18.Supplement_1.S79.
[72] T. Chalder *et al.*, “Development of a fatigue scale,” *J. Psychosom. Res.*, vol. 37, no. 2, pp. 147–153, Feb. 1993, doi: 10.1016/0022-3999(93)90081-P.
[73] L. B. Krupp, N. G. LaRocca, J. Muir-Nash, and A. D. Steinberg, “The Fatigue Severity Scale: Application to Patients With Multiple Sclerosis and Systemic Lupus Erythematosus,” *Arch. Neurol.*, vol. 46, no. 10, pp. 1121–1123, Oct. 1989, doi: 10.1001/archneur.1989.00520460115022.
[74] B. B. Tack, “Dimensions and correlates of fatigue in older adults with rheumatoid arthritis.,” Accessed: Jun. 24, 2020. [Online]. Available: https://elibrary.ru/item.asp?id=5833962.
[75] E. M. A. Smets, B. Garssen, B. Bonke, and J. C. J. M. De Haes, “The multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue,” *J. Psychosom. Res.*, vol. 39, no. 3, pp. 315–325, Apr. 1995, doi: 10.1016/0022-3999(94)00125-O.
[76] J. E. Ware Jr and C. D. Sherbourne, “The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection,” *Med. Care*, pp. 473–483, 1992.
[77] J.-S. Lai *et al.*, “How Item Banks and Their Application Can Influence Measurement Practice in Rehabilitation Medicine: A PROMIS Fatigue Item Bank Example,” *Arch. Phys. Med. Rehabil.*, vol. 92, no. 10, pp. S20–S27, Oct. 2011, doi: 10.1016/j.apmr.2010.08.033.
[78] D. McNair, M. Lorr, and L. Droppleman, “POMS manual for the Profile of Mood States,” *Educ. Ind. Test. Serv.*, 1971.
[79] K. A. Lee, G. Hicks, and G. Nino-Murcia, “Validity and reliability of a scale to assess fatigue,” *Psychiatry Res.*, vol. 36, no. 3, pp. 291–298, Mar. 1991, doi: 10.1016/0165-1781(91)90027-M.
[80] D. U. Junghaenel, C. Christodoulou, J.-S. Lai, and A. A. Stone, “Demographic Correlates of Fatigue in the US General Population: Results from the Patient-Reported Outcomes Measurement Information System (PROMIS) Initiative,” *J. Psychosom. Res.*, vol. 71, no. 3, pp. 117–123, Sep. 2011, doi: 10.1016/j.jpsychores.2011.04.007.
[81] Z. B. Alemohammad and K. Sadeghniiat-Haghighi, “Risk of Fatigue at Work,” in *Fatigue Management: Principles and Practices for Improving Workplace Safety*, A. Sharafkhaneh and M. Hirshkowitz, Eds. New York, NY: Springer, 2018, pp. 181–191.
[82] A. Sharafkhaneh and M. Hirshkowitz, Eds., *Fatigue Management: Principles and Practices for Improving Workplace Safety*. New York, NY: Springer New York, 2018.
[83] S. Xianglong, Z. Hu, F. Shumin, and L. Zhenning, “Bus drivers’ mood states and reaction abilities at high temperatures,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 59, pp. 436–444, Nov. 2018, doi: 10.1016/j.trf.2018.09.022.
[84] S. A. Mitchell, “Cancer-Related Fatigue: State of the Science,” *PM&R*, vol. 2, no. 5, pp. 364–383, May 2010, doi: 10.1016/j.pmrj.2010.03.024.
[85] E. J. M. Pearson, M. E. Morris, M. di Stefano, and C. E. McKinstry, “Interventions for cancer-related fatigue: a scoping review,” *Eur. J. Cancer Care (Engl.)*, vol. 27, no. 1, p. e12516, Jan. 2018, doi: 10.1111/ecc.12516.
[86] R. Cohen, C. Bavishi, S. Haider, J. Thankachen, and A. Rozanski, “Meta-Analysis of Relation of Vital Exhaustion to Cardiovascular Disease Events,” *Am. J. Cardiol.*, vol. 119, no. 8, pp. 1211–1216, Apr. 2017, doi: 10.1016/j.amjcard.2017.01.009.
[87] A. Nagy *et al.*, “Association of Exercise Capacity with Physical Functionality and Various Aspects of Fatigue in Patients with Coronary Artery Disease,” *Behav. Med.*, vol. 44, no. 1, pp. 28–35, Jan. 2018,

doi: 10.1080/08964289.2016.1189395.
[88] A. Lewko, P. Bidgood, A. Jewell, and R. Garrod, “A Comprehensive Literature Review of COPD-Related Fatigue,” *Curr. Respir. Med. Rev.*, vol. 8, no. 5, pp. 370–382, Nov. 2012, doi: 10.2174/157339812803832476.
[89] Ø. Jensen, T. Bernklev, and L.-P. Jelsness-Jørgensen, “Fatigue in type 1 diabetes: A systematic review of Observational studies,” *Diabetes Res. Clin. Pract.*, vol. 123, pp. 63–74, Jan. 2017, doi: 10.1016/j.diabres.2016.11.002.
[90] A. Vincent, R. P. Benzo, M. O. Whipple, S. J. McAllister, P. J. Erwin, and L. N. Saligan, “Beyond pain in fibromyalgia: insights into the symptom of fatigue,” *Arthritis Res. Ther.*, vol. 15, no. 6, p. 221, Nov. 2013, doi: 10.1186/ar4395.
[91] W. Czuber-Dochan, E. Ream, and C. Norton, “Review article: description and management of fatigue in inflammatory bowel disease,” *Aliment. Pharmacol. Ther.*, vol. 37, no. 5, pp. 505–516, 2013, doi: 10.1111/apt.12205.
[92] C. J. Han and G. S. Yang, “Fatigue in Irritable Bowel Syndrome: A Systematic Review and Meta-analysis of Pooled Frequency and Severity of Fatigue,” *Asian Nurs. Res.*, vol. 10, no. 1, pp. 1–10, Mar. 2016, doi: 10.1016/j.anr.2016.01.003.
[93] A. Nocerino, A. Nguyen, M. Agrawal, A. Mone, K. Lakhani, and A. Swaminath, “Fatigue in Inflammatory Bowel Diseases: Etiologies and Management,” *Adv. Ther.*, vol. 37, no. 1, pp. 97–112, Jan. 2020, doi: 10.1007/s12325-019-01151-w.
[94] I. Moldovan *et al.*, “Pain and depression predict self-reported fatigue/energy in lupus,” *Lupus*, vol. 22, no. 7, pp. 684–689, Jun. 2013, doi: 10.1177/0961203313486948.
[95] H. Ghanean, A. K. Ceniti, and S. H. Kennedy, “Fatigue in Patients with Major Depressive Disorder: Prevalence, Burden and Pharmacological Approaches to Management,” *CNS Drugs*, vol. 32, no. 1, pp. 65–74, Jan. 2018, doi: 10.1007/s40263-018-0490-z.
[96] L. Charvet, D. Serafin, and L. B. Krupp, “Fatigue in multiple sclerosis,” *Fatigue Biomed. Health Behav.*, vol. 2, no. 1, pp. 3–13, Jan. 2014, doi: 10.1080/21641846.2013.843812.
[97] A. E. Latimer-Cheung *et al.*, “Effects of Exercise Training on Fitness, Mobility, Fatigue, and Health-Related Quality of Life Among Adults With Multiple Sclerosis: A Systematic Review to Inform Guideline Development,” *Arch. Phys. Med. Rehabil.*, vol. 94, no. 9, pp. 1800-1828.e3, Sep. 2013, doi: 10.1016/j.apmr.2013.04.020.
[98] P. Newland, A. Starkweather, and M. Sorenson, “Central fatigue in multiple sclerosis: a review of the literature,” *J. Spinal Cord Med.*, vol. 39, no. 4, pp. 386–399, Jul. 2016, doi: 10.1080/10790268.2016.1168587.
[99] I. M. Skoie *et al.*, “Fatigue in psoriasis: a controlled study,” *Br. J. Dermatol.*, vol. 177, no. 2, pp. 505–512, 2017.
[100] A. A. Weinstein *et al.*, “Effect of aerobic exercise training on fatigue and physical activity in patients with pulmonary arterial hypertension,” *Respir. Med.*, vol. 107, no. 5, pp. 778–784, May 2013, doi: 10.1016/j.rmed.2013.02.006.
[101] M. Artom, R. Moss-Morris, F. Caskey, and J. Chilcot, “Fatigue in advanced kidney disease,” *Kidney Int.*, vol. 86, no. 3, pp. 497–505, Sep. 2014, doi: 10.1038/ki.2014.86.
[102] S. Stebbings and G. J. Treharne, “Fatigue in rheumatic disease: an overview,” *Int. J. Clin. Rheumatol.*, vol. 5, no. 4, pp. 487–502, Aug. 2010, doi: 10.2217/ijr.10.30.
[103] R. C. do Espírito Santo *et al.*, “Neuromuscular fatigue is weakly associated with perception of fatigue and function in patients with rheumatoid arthritis,” *Rheumatol. Int.*, vol. 38, no. 3, pp. 415–423, Mar. 2018, doi: 10.1007/s00296-017-3894-z.
[104] S. Bailes *et al.*, “Fatigue: The forgotten symptom of sleep apnea,” *J. Psychosom. Res.*, vol. 70, no. 4, pp. 346–354, Apr. 2011, doi: 10.1016/j.jpsychores.2010.09.009.
[105] R. Aarnes, J. Stubberud, and A. Lerdal, “A literature review of factors associated with fatigue

after stroke and a proposal for a framework for clinical utility," *Neuropsychol. Rehabil.*, vol. 0, no. 0, pp. 1–28, Mar. 2019, doi: 10.1080/09602011.2019.1589530.
[106] F. Duncan, S. Wu, and G. E. Mead, "Frequency and natural history of fatigue after stroke: A systematic review of longitudinal studies," *J. Psychosom. Res.*, vol. 73, no. 1, pp. 18–27, Jul. 2012, doi: 10.1016/j.jpsychores.2012.04.001.
[107] P. Newland, B. Oliver, J. M. Newland, and F. P. Thomas, "Testing Feasibility of a Mobile Application to Monitor Fatigue in People With Multiple Sclerosis," *J. Neurosci. Nurs.*, vol. 51, no. 6, pp. 331–334, Dec. 2019, doi: 10.1097/JNN.0000000000000479.
[108] R. Maguire, L. McCann, M. Miller, and N. Kearney, "Nurse's perceptions and experiences of using of a mobile-phone-based Advanced Symptom Management System (ASyMS©) to monitor and manage chemotherapy-related toxicity," *Eur. J. Oncol. Nurs.*, vol. 12, no. 4, pp. 380–386, Sep. 2008, doi: 10.1016/j.ejon.2008.04.007.
[109] J. Nicholas, M. E. Larsen, J. Proudfoot, and H. Christensen, "Mobile Apps for Bipolar Disorder: A Systematic Review of Features and Content Quality," *J. Med. Internet Res.*, vol. 17, no. 8, p. e198, 2015, doi: 10.2196/jmir.4581.
[110] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological Momentary Assessment," *Annu. Rev. Clin. Psychol.*, vol. 4, no. 1, pp. 1–32, 2008, doi: 10.1146/annurev.clinpsy.3.022806.091415.
[111] R. Larson and M. Csikszentmihalyi, "The Experience Sampling Method," in *Flow and the Foundations of Positive Psychology*, Dordrecht: Springer Netherlands, 2014, pp. 21–34.
[112] N. van Berkel, D. Ferreira, and V. Kostakos, "The Experience Sampling Method on Mobile Devices," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–40, Jan. 2018, doi: 10.1145/3123988.
[113] E. Guérin, M. S. Fortier, and T. Williams, "'I just NEED to move…': examining women's passion for physical activity and its relationship with daily affect and vitality," *Psychol. Well- Theory Res. Pract.*, vol. 3, no. 1, p. 4, Nov. 2013, doi: 10.1186/2211-1522-3-4.
[114] H. Zacher, H. A. Brailsford, and S. L. Parker, "Micro-breaks matter: A diary study on the effects of energy management strategies on occupational well-being," *J. Vocat. Behav.*, vol. 85, no. 3, pp. 287–297, Dec. 2014, doi: 10.1016/j.jvb.2014.08.005.
[115] H. Rhee and S. Kim, "Effects of breaks on regaining vitality at work: An empirical comparison of 'conventional' and 'smart phone' breaks," *Comput. Hum. Behav.*, vol. 57, pp. 160–167, Apr. 2016, doi: 10.1016/j.chb.2015.11.056.
[116] S. L. Murphy and D. M. Smith, "Ecological Measurement of Fatigue and Fatigability in Older Adults With Osteoarthritis," *J. Gerontol. Ser. A*, vol. 65A, no. 2, pp. 184–189, Feb. 2010, doi: 10.1093/gerona/glp137.
[117] E. D. Hacker and C. E. Ferrans, "Ecological Momentary Assessment of Fatigue in Patients Receiving Intensive Cancer Therapy," *J. Pain Symptom Manage.*, vol. 33, no. 3, pp. 267–275, Mar. 2007, doi: 10.1016/j.jpainsymman.2006.08.007.
[118] J. E. Bower, "Cancer-related fatigue: Mechanisms, risk factors, and treatments," *Nat. Rev. Clin. Oncol.*, vol. 11, no. 10, pp. 597–609, Oct. 2014, doi: 10.1038/nrclinonc.2014.127.
[119] D. van der Linden, M. Frese, and T. F. Meijman, "Mental fatigue and the control of cognitive processes: effects on perseveration and planning," *Acta Psychol. (Amst.)*, vol. 113, no. 1, pp. 45–65, May 2003, doi: 10.1016/S0001-6918(02)00150-6.
[120] P. L. Ackerman, "100 years without resting," in *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications*, Washington, DC, US: American Psychological Association, 2011, pp. 11–43.
[121] H. R. Lieberman, "Cognitive methods for assessing mental energy," *Nutr. Neurosci.*, vol. 10, no. 5–6, pp. 229–242, Oct. 2007, doi: 10.1080/10284150701722273.
[122] M. Basner and D. F. Dinges, "Maximizing Sensitivity of the Psychomotor Vigilance Test (PVT) to Sleep Loss," *Sleep*, vol. 34, no. 5, pp. 581–591, May 2011, doi: 10.1093/sleep/34.5.581.

[123] L. Holsbeeke, M. Ketelaar, M. M. Schoemaker, and J. W. Gorter, “Capacity, Capability, and Performance: Different Constructs or Three of a Kind?,” *Arch. Phys. Med. Rehabil.*, vol. 90, no. 5, pp. 849–855, May 2009, doi: 10.1016/j.apmr.2008.11.015.
[124] World Health Organization, “World report on ageing and health.,” *World Health Organ.*, 2015.
[125] P. L. Enright, “The Six-Minute Walk Test,” *Respir. CARE*, vol. 48, no. 8, p. 3, 2003.
[126] R. Rudick *et al.*, “Recommendations from the national multiple sclerosis society clinical outcomes assessment task force,” *Ann. Neurol.*, vol. 42, no. 3, pp. 379–382, 1997, doi: 10.1002/ana.410420318.
[127] D. Podsiadlo and S. Richardson, “The Timed ‘Up & Go’: A Test of Basic Functional Mobility for Frail Elderly Persons,” *J. Am. Geriatr. Soc.*, vol. 39, no. 2, pp. 142–148, 1991, doi: 10.1111/j.1532-5415.1991.tb01616.x.
[128] A. Boolani, P. J. O’Connor, J. Reid, S. Ma, and S. Mondal, “Predictors of feelings of energy differ from predictors of fatigue,” *Fatigue Biomed. Health Behav.*, vol. 7, no. 1, pp. 12–28, Jan. 2019, doi: 10.1080/21641846.2018.1558733.
[129] C. Roldán Jiménez, P. Bennett, A. Ortiz García, and A. I. Cuesta Vargas, “Fatigue Detection during Sit-To-Stand Test Based on Surface Electromyography and Acceleration: A Case Study,” *Sensors*, vol. 19, no. 19, Art. no. 19, Jan. 2019, doi: 10.3390/s19194202.
[130] J. E. Muñoz, E. R. Gouveia, M. S. Cameirão, and S. B. i Badia, “PhysioLab - a multivariate physiological computing toolbox for ECG, EMG and EDA signals: a case of study of cardiorespiratory fitness assessment in the elderly population,” *Multimed. Tools Appl.*, vol. 77, no. 9, pp. 11521–11546, May 2018, doi: 10.1007/s11042-017-5069-z.
[131] B. J. Nicklas *et al.*, “Effects of Caloric Restriction on Cardiorespiratory Fitness, Fatigue, and Disability Responses to Aerobic Exercise in Older Adults With Obesity: A Randomized Controlled Trial,” *J. Gerontol. A. Biol. Sci. Med. Sci.*, vol. 74, no. 7, pp. 1084–1090, Jun. 2019, doi: 10.1093/gerona/gly159.
[132] A. Sehle, M. Vieten, A. Mündermann, and C. Dettmers, “Difference in Motor Fatigue between Patients with Stroke and Patients with Multiple Sclerosis: A Pilot Study,” *Front. Neurol.*, vol. 5, 2014, doi: 10.3389/fneur.2014.00279.
[133] Å. Dedering, Å. Gnospelius, and B. Elfving, “Reliability of measurements of endurance time, electromyographic fatigue and recovery, and associations to activity limitations, in patients with lumbar disc herniation,” *Physiother. Res. Int.*, vol. 15, no. 4, pp. 189–198, 2010, doi: 10.1002/pri.457.
[134] J. Schmitt, N. Lindner, M. Reuss‐Borst, H.-C. Holmberg, and B. Sperlich, “A 3-week multimodal intervention involving high-intensity interval training in female cancer survivors: a randomized controlled trial,” *Physiol. Rep.*, vol. 4, no. 3, p. e12693, 2016, doi: 10.14814/phy2.12693.
[135] M. F. Folstein, L. N. Robins, and J. E. Helzer, “The Mini-Mental State Examination,” *Arch. Gen. Psychiatry*, vol. 40, no. 7, pp. 812–812, Jul. 1983, doi: 10.1001/archpsyc.1983.01790060110016.
[136] U. S. Army, “Army individual test battery,” *Man. Dir. Scoring*, 1944.
[137] H. J. Hannay and H. S. Levin, “Selective reminding test: An examination of the equivalence of four forms,” *J. Clin. Exp. Neuropsychol.*, vol. 7, no. 3, pp. 251–263, Jun. 1985, doi: 10.1080/01688638508401258.
[138] J. Barbizet and E. Cany, “Clinical and psychometrical study of a patient with memory disturbances,” *Int. J. Neurol.*, vol. 7, no. 1, pp. 44–54, 1967.
[139] S. M. Rao, T. A. Hammeke, M. P. McQuillen, B. O. Khatri, and D. Lloyd, “Memory Disturbance in Chronic Progressive Multiple Sclerosis,” *Arch. Neurol.*, vol. 41, no. 6, pp. 625–631, Jun. 1984, doi: 10.1001/archneur.1984.04210080033010.
[140] A. Smith, *Symbol digit modalities test*. Western Psychological Services Los Angeles, 1973.
[141] D. M. A. Gronwall, “Paced Auditory Serial-Addition Task: A Measure of Recovery from

Concussion," *Percept. Mot. Skills*, vol. 44, no. 2, pp. 367–373, Apr. 1977, doi: 10.2466/pms.1977.44.2.367.
[142] A. Barr and J. Brandt, "Word-list generation deficits in dementia," *J. Clin. Exp. Neuropsychol.*, vol. 18, no. 6, pp. 810–822, Dec. 1996, doi: 10.1080/01688639608408304.
[143] A. Rey, "L'examen clinique en psychologie.," 1958.
[144] P. A. Obrist, R. A. Webb, and J. R. Sutterer, "Heart rate and somatic changes during aversive conditioning and a simple reaction time task," *Psychophysiology*, vol. 5, no. 6, pp. 696–723, 1969.
[145] C. T. Bever Jr, L. Grattan, H. S. Panitch, and K. P. Johnson, "The brief repeatable battery of neuropsychological tests for multiple sclerosis: a preliminary serial study," *Mult. Scler. J.*, vol. 1, no. 3, pp. 165–169, 1995.
[146] K. C. Smolders and Y. A. de Kort, "Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal," *J. Environ. Psychol.*, vol. 39, pp. 77–91, 2014.
[147] J. I. Westbrook, M. Z. Raban, S. R. Walter, and H. Douglas, "Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: a prospective, direct observation study," *BMJ Qual. Saf.*, vol. 27, no. 8, pp. 655–663, 2018.
[148] B. M. Hartzler, "Fatigue on the flight deck: the consequences of sleep loss and the benefits of napping," *Accid. Anal. Prev.*, vol. 62, pp. 309–318, 2014.
[149] M. Staniute, A. Bunevicius, J. Brozaitiene, and R. Bunevicius, "Relationship of health-related quality of life with fatigue and exercise capacity in patients with coronary artery disease," *Eur. J. Cardiovasc. Nurs.*, vol. 13, no. 4, pp. 338–344, Aug. 2014, doi: 10.1177/1474515113496942.
[150] A. Pimenta, D. Carneiro, P. Novais, and J. Neves, "Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns," in *International Conference on Hybrid Artificial Intelligence Systems*, 2013, pp. 222–231.
[151] K. Foti, D. Xanthopoulou, S. Papagiannidis, and K. Kafetsios, "The Role of Tweet-Related Emotion on the Exhaustion–Recovery from Work Relationship," in *Conference on e-Business, e-Services and e-Society*, 2019, pp. 380–391.
[152] S. Dikmen, J. Machamer, N. Temkin, and A. McLean, "Neuropsychological recovery in patients with moderate to severe head injury: 2 year follow-up," *J. Clin. Exp. Neuropsychol.*, vol. 12, no. 4, pp. 507–519, 1990.
[153] J. L. Ponsford, J. H. Olver, and C. Curran, "A profile of outcome: 2 years after traumatic brain injury," *Brain Inj.*, vol. 9, no. 1, pp. 1–10, 1995.
[154] A. H. Van Zomeren and W. Van den Burg, "Residual complaints of patients two years after severe head injury.," *J. Neurol. Neurosurg. Psychiatry*, vol. 48, no. 1, pp. 21–28, 1985.
[155] E. Price, G. Moore, L. Galway, and M. Linden, "Towards a mobile assistive technology for monitoring and assessing cognitive fatigue in individuals with acquired brain injury," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 1487–1491.
[156] Y. Tran, N. Wijesuriya, M. Tarvainen, P. Karjalainen, and A. Craig, "The relationship between spectral changes in heart rate variability and fatigue," *J. Psychophysiol.*, vol. 23, no. 3, pp. 143–151, 2009.
[157] S. C. Segerstrom and L. S. Nes, "Heart rate variability reflects self-regulatory strength, effort, and fatigue," *Psychol. Sci.*, vol. 18, no. 3, pp. 275–281, 2007.
[158] A. Samel, H.-M. Wegmann, and M. Vejvoda, "Aircrew fatigue in long-haul operations," *Accid. Anal. Prev.*, vol. 29, no. 4, pp. 439–452, 1997.
[159] L. I. Goldman, M. T. McDonough, and G. P. Rosemond, "Stresses affecting surgical performance and learning: I. Correlation of heart rate, electrocardiogram, and operation simultaneously recorded on videotapes," *J. Surg. Res.*, vol. 12, no. 2, pp. 83–86, 1972.

[160] C. Chen, K. Li, Q. Wu, H. Wang, Z. Qian, and G. Sudlow, “EEG-based detection and evaluation of fatigue caused by watching 3DTV,” *Displays*, vol. 34, no. 2, pp. 81–88, 2013.
[161] L. Holdsworth and T. Evens, “Is heart rate variability a useful marker of stress and fatigue in emergency and pre-hospital clinicians?–A systematic review,” *Resuscitation*, vol. 118, p. e95, 2017.
[162] J. S. Stamler, M. E. Goldman, J. Gomes, D. Matza, and S. F. Horowitz, “The effect of stress and fatigue on cardiac rhythm in medical interns,” *J. Electrocardiol.*, vol. 25, no. 4, pp. 333–338, 1992.
[163] J. Heemskerk *et al.*, “Relax, it’s just laparoscopy! A prospective randomized trial on heart rate variability of the surgeon in robot-assisted versus conventional laparoscopic cholecystectomy,” *Dig. Surg.*, vol. 31, no. 3, pp. 225–232, 2014.
[164] S.-J. Jung, H.-S. Shin, and W.-Y. Chung, “Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel,” *IET Intell. Transp. Syst.*, vol. 8, no. 1, pp. 43–50, 2014.
[165] L. G. Díaz-Toral, T. R. Banderas-Dorantes, and J. F. Rivas-Vilchis, “Impact of electroacupuncture treatment on quality of life and heart rate variability in fibromyalgia patients,” *J. Evid.-Based Complement. Altern. Med.*, vol. 22, no. 2, pp. 216–222, 2017.
[166] J. DeLuca, “19 Fatigue: Its Definition, Its Study, and Its Future,” *Fatigue Window Brain*, p. 319, 2005.
[167] J. A. Stern, D. Boyer, and D. Schroeder, “Blink rate: a possible measure of fatigue,” *Hum. Factors*, vol. 36, no. 2, pp. 285–297, 1994.
[168] S. Sirois and J. Brisson, “Pupillometry,” *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 5, no. 6, pp. 679–692, 2014.
[169] T. Eggert *et al.*, “The pupillographic sleepiness test in adults: Effect of age, gender, and time of day on pupillometric variables,” *Am. J. Hum. Biol.*, vol. 24, no. 6, pp. 820–828, 2012.
[170] O. Lowenstein, R. Feinberg, and I. E. Loewenfeld, “Pupillary movements during acute and chronic fatigue: A new test for the objective evaluation of tiredness,” *Invest. Ophthalmol. Vis. Sci.*, vol. 2, no. 2, pp. 138–157, 1963.
[171] K. Hirvonen, S. Puttonen, K. Gould, J. Korpela, V. F. Koefoed, and K. Müller, “Improving the saccade peak velocity measurement for detecting fatigue,” *J. Neurosci. Methods*, vol. 187, no. 2, pp. 199–206, 2010.
[172] C.-W. You *et al.*, “CarSafe: a driver safety app that detects dangerous driving behavior using dual-cameras on smartphones,” in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 671–672.
[173] Y.-F. Zhang, X.-Y. Gao, J.-Y. Zhu, W.-L. Zheng, and B.-L. Lu, “A novel approach to driving fatigue detection using forehead EOG,” in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2015, pp. 707–710.
[174] G. Morris and M. Maes, “Myalgic encephalomyelitis/chronic fatigue syndrome and encephalomyelitis disseminata/multiple sclerosis show remarkable levels of similarity in phenomenology and neuroimmune characteristics,” *BMC Med.*, vol. 11, no. 1, pp. 1–23, 2013.
[175] M. M. Lorist, E. Bezdan, M. ten Caat, M. M. Span, J. B. Roerdink, and N. M. Maurits, “The influence of mental fatigue and motivation on neural network dynamics; an EEG coherence study,” *Brain Res.*, vol. 1270, pp. 95–106, 2009.
[176] Y. Sheng, K. Huang, J. Zou, L. Wang, and P. Wei, “Exploring the Relationship between Neural Mechanism and Detection in Mental Fatigue by Genetic Algorithm and Hierarchical Clustering,” in *2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB)*, 2019, pp. 120–124.
[177] A. Craig, Y. Tran, N. Wijesuriya, and H. Nguyen, “Regional brain wave activity changes associated with fatigue,” *Psychophysiology*, vol. 49, no. 4, pp. 574–582, 2012.
[178] C. Zhao, C. Zheng, M. Zhao, Y. Tu, and J. Liu, “Multivariate autoregressive models and kernel

learning algorithms for classifying driving mental fatigue based on electroencephalographic," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 1859–1865, 2011.
[179] K. Kahol, M. Smith, J. Brandenberger, A. Ashby, and J. J. Ferrara, "Impact of fatigue on neurophysiologic measures of surgical residents," *J. Am. Coll. Surg.*, vol. 213, no. 1, pp. 29–34, 2011.
[180] K. A. Guru, S. B. Shafiei, A. Khan, A. A. Hussein, M. Sharif, and E. T. Esfahani, "Understanding cognitive performance during robot-assisted surgery," *Urology*, vol. 86, no. 4, pp. 751–757, 2015.
[181] A. E. Gold, K. M. MacLeod, I. J. Deary, and B. M. Frier, "Hypoglycemia-induced cognitive dysfunction in diabetes mellitus: effect of hypoglycemia unawareness," *Physiol. Behav.*, vol. 58, no. 3, pp. 501–511, 1995.
[182] E. A. Newsholme, E. Blomstrand, and B. Ekblom, "Physical and mental fatigue: metabolic mechanisms and importance of plasma amino acids," *Br. Med. Bull.*, vol. 48, no. 3, pp. 477–495, 1992.
[183] E. M. Olsson, W. T. Roth, and L. Melin, "Psychophysiological characteristics of women suffering from stress-related fatigue," *Stress Health J. Int. Soc. Investig. Stress*, vol. 26, no. 2, pp. 113–126, 2010.
[184] U. Kuebler *et al.*, "Norepinephrine infusion with and without alpha-adrenergic blockade by phentolamine increases salivary alpha amylase in healthy men," *Psychoneuroendocrinology*, vol. 49, pp. 290–298, 2014.
[185] J. F. Hopstaken, D. van der Linden, A. B. Bakker, and M. A. Kompier, "The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics," *Biol. Psychol.*, vol. 110, pp. 100–106, 2015.
[186] R. O. Van Heukelom, J. B. Prins, M. G. Smits, and G. Bleijenberg, "Influence of melatonin on fatigue severity in patients with chronic fatigue syndrome and late melatonin secretion," *Eur. J. Neurol.*, vol. 13, no. 1, pp. 55–60, 2006.
[187] K. R. Merikangas *et al.*, "Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder," *JAMA Psychiatry*, vol. 76, no. 2, pp. 190–198, 2019.
[188] K. E. Bloch, "Polysomnography: a systematic review," *Technol. Health Care*, vol. 5, no. 4, pp. 285–305, 1997.
[189] A. Sadeh, P. J. Hauri, D. F. Kripke, and P. Lavie, "The role of actigraphy in the evaluation of sleep disorders," *Sleep*, vol. 18, no. 4, pp. 288–302, 1995.
[190] K. Wac, "From quantified self to quality of life," in *Digital Health*, Springer, 2018, pp. 83–108.
[191] D. Rizzo *et al.*, "The role of fatigue and sleepiness in drivers with obstructive sleep apnea," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 62, pp. 796–804, 2019.
[192] S. A. Fernandes-Junior, F. S. Ruiz, L. S. Antonietti, S. Tufik, and M. Túlio de Mello, "Sleep, fatigue and quality of life: A comparative analysis among night shift workers with and without children," *PloS One*, vol. 11, no. 7, p. e0158580, 2016.
[193] R. Campbell, E. Tobback, L. Delesie, D. Vogelaers, A. Mariman, and M. Vansteenkiste, "Basic psychological need experiences, fatigue, and sleep in individuals with unexplained chronic fatigue," *Stress Health*, vol. 33, no. 5, pp. 645–655, 2017.
[194] C. Maher *et al.*, "Fatigue is a major issue for children and adolescents with physical disabilities," *Dev. Med. Child Neurol.*, vol. 57, no. 8, pp. 742–747, 2015.
[195] H. P. Attarian, K. M. Brown, S. P. Duntley, J. D. Carter, and A. H. Cross, "The Relationship of Sleep Disturbances and Fatigue in Multiple Sclerosis," *Arch. Neurol.*, vol. 61, no. 4, pp. 525–528, Apr. 2004, doi: 10.1001/archneur.61.4.525.
[196] H. Kaynak *et al.*, "Fatigue and sleep disturbance in multiple sclerosis," *Eur. J. Neurol.*, vol. 13, no. 12, pp. 1333–1339, 2006.

[197] C. Veauthier *et al.*, “Fatigue in multiple sclerosis is closely related to sleep disorders: a polysomnographic cross-sectional study,” *Mult. Scler. J.*, vol. 17, no. 5, pp. 613–622, 2011.
[198] M. Kaminska *et al.*, “Obstructive sleep apnea is associated with fatigue in multiple sclerosis,” *Mult. Scler. J.*, vol. 18, no. 8, pp. 1159–1169, 2012.
[199] A. B.-G. Shitrit *et al.*, “Sleep disturbances can be prospectively observed in patients with an inactive inflammatory bowel disease,” *Dig. Dis. Sci.*, vol. 63, no. 11, pp. 2992–2997, 2018.
[200] J.-L. Sun and C.-C. Lin, “Relationships among daytime napping and fatigue, sleep quality, and quality of life in cancer patients,” *Cancer Nurs.*, vol. 39, no. 5, pp. 383–392, 2016.
[201] S. Ancoli-Israel *et al.*, “Sleep, fatigue, depression, and circadian activity rhythms in women with breast cancer before and after treatment: a 1-year longitudinal study,” *Support. Care Cancer*, vol. 22, no. 9, pp. 2535–2545, Sep. 2014, doi: 10.1007/s00520-014-2204-5.
[202] E. B. Holliday, N. F. Dieckmann, T. L. McDonald, A. Y. Hung, C. R. Thomas Jr, and L. J. Wood, “Relationship between fatigue, sleep quality and inflammatory cytokines during external beam radiation therapy for prostate cancer: A prospective study,” *Radiother. Oncol.*, vol. 118, no. 1, pp. 105–111, 2016.
[203] T. Cambras, J. Castro-Marrero, M. C. Zaragoza, A. Díez-Noguera, and J. Alegre, “Circadian rhythm abnormalities and autonomic dysfunction in patients with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis,” *PloS One*, vol. 13, no. 6, p. e0198106, 2018.
[204] B. J. Nicklas, D. P. Beavers, S. L. Mihalko, G. D. Miller, R. F. Loeser, and S. P. Messier, “Relationship of objectively-measured habitual physical activity to chronic inflammation and fatigue in middle-aged and older adults,” *J. Gerontol. Ser. Biomed. Sci. Med. Sci.*, vol. 71, no. 11, pp. 1437–1443, 2016.
[205] M. Nilsson, C. Arving, I. Thormodsen, J. Assmus, S. Berntsen, and K. Nordin, “Moderate-to-vigorous intensity physical activity is associated with modified fatigue during and after cancer treatment,” *Support. Care Cancer*, pp. 1–8, 2019.
[206] D. Vancampfort *et al.*, “Higher cardio-respiratory fitness is associated with increased mental and physical quality of life in people with bipolar disorder: A controlled pilot study,” *Psychiatry Res.*, vol. 256, pp. 219–224, 2017.
[207] A. Sheshadri, P. Kittiskulnam, and K. L. Johansen, “Higher physical activity is associated with less fatigue and insomnia among patients on hemodialysis,” *Kidney Int. Rep.*, vol. 4, no. 2, pp. 285–292, 2019.
[208] T. Qazi and F. A. Farraye, “Sleep and inflammatory bowel disease: an important bi-directional relationship,” *Inflamm. Bowel Dis.*, vol. 25, no. 5, pp. 843–852, 2019.
[209] M. A. Sofia, M. Andersen, and D. T. Rubin, “Sleep dysfunction and inflammatory bowel disease,” in *Sleep Effect on Gastrointestinal Health and Disease: Translational Opportunities for Promoting Health and Optimizing Disease Management*, Nova Science Publishers, Inc., 2018, pp. 177–199.
[210] L. Abbott, “Energy Through Motion\copyright: An activity intervention for cancer-related fatigue in an ambulatory infusion center,” *Clin. J. Oncol. Nurs.*, vol. 21, no. 5, p. 618, 2017.
[211] D. Ren, Y. An, and Z. Li, “The Discriminative Model of Mental Fatigue Based on Comprehensive Parameter Analysis,” in *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, 2019, pp. 1–6.
[212] M. R. Smith, R. Chai, H. T. Nguyen, S. M. Marcora, and A. J. Coutts, “Comparing the effects of three cognitive tasks on indicators of mental fatigue,” *J. Psychol.*, vol. 153, no. 8, pp. 759–783, 2019.
[213] D. M. Brown and S. R. Bray, “Effects of mental fatigue on exercise intentions and behavior,” *Ann. Behav. Med.*, vol. 53, no. 5, pp. 405–414, 2019.
[214] J. L. Kanitz *et al.*, “The impact of eurythmy therapy on fatigue in healthy adults—a controlled

trial," *Eur. J. Integr. Med.*, vol. 4, no. 3, pp. e289–e297, 2012.
[215] K. C. Smolders, Y. A. De Kort, and P. J. M. Cluitmans, "A higher illuminance induces alertness even during office hours: findings on subjective measures, task performance and heart rate measures," *Physiol. Behav.*, vol. 107, no. 1, pp. 7–16, 2012.
[216] S. Oriyama, Y. Miyakoshi, and T. Kobayashi, "Effects of two 15-min naps on the subjective sleepiness, fatigue and heart rate variability of night shift nurses," *Ind. Health*, 2013.
[217] H. Singh *et al.*, "Robotic surgery improves technical performance and enhances prefrontal activation during high temporal demand," *Ann. Biomed. Eng.*, vol. 46, no. 10, pp. 1621–1636, 2018.
[218] R. K. Dishman, N. J. Thom, T. W. Puetz, P. J. O'Connor, and B. A. Clementz, "Effects of cycling exercise on vigor, fatigue, and electroencephalographic activity among young adults who report persistent fatigue," *Psychophysiology*, vol. 47, no. 6, pp. 1066–1074, 2010.
[219] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 39–58, 2008.
[220] T. J. Braley and R. D. Chervin, "Fatigue in multiple sclerosis: mechanisms, evaluation, and treatment," *Sleep*, vol. 33, no. 8, pp. 1061–1067, 2010.
[221] A. Adinolfi, "Assessment and Treatment of HIV-Related Fatigue," *J. Assoc. Nurses AIDS Care*, vol. 12, pp. 29–34, Sep. 2001, doi: 10.1016/S1055-3290(06)60155-6.
[222] K. Wac, "Quality of Life Technologies," *Encycl. Behav. Med. Springer N. Y. NY USA*, 2020.
[223] M. Swan, "The quantified self: Fundamental disruption in big data science and biological discovery," *Big Data*, vol. 1, no. 2, pp. 85–99, 2013.
[224] A. Neuringer, "Self-experimentation: A call for change," *Behaviorism*, vol. 9, no. 1, pp. 79–94, 1981.
[225] S. Jones, "Health & Fitness Wearables," 2014. https://www.juniperresearch.com/researchstore/devices-technology/digital-health-wearables-research-report (accessed Jun. 19, 2020).
[226] L. Piwek, D. A. Ellis, S. Andrews, and A. Joinson, "The rise of consumer health wearables: promises and barriers," *PLoS Med.*, vol. 13, no. 2, p. e1001953, 2016.
[227] E. A. Vogels, "About one-in-five Americans use a smart watch or fitness tracker," *Wash. DC Pew Res. Cent.*, 2019.
[228] J. Ubrani, R. Llamas, and M. Shirer, "IDC forecasts sustained double-digit growth for wearable devices led by steady adoption of smart-watches," *International Data Corporation*, 2020. https://www.idc.com/getdoc.jsp?containerId=p (accessed Jun. 15, 2020).
[229] V. Estrada-Galiñanes and K. Wac, "Collecting, exploring and sharing personal data: Why, how and where," *Data Sci.*, no. Preprint, pp. 1–28.
[230] K. Stansberry, J. Anderson, and L. Rainie, "Experts Optimistic About the Next 50 Years of Digital Life," *Pew Research Center: Internet, Science & Tech*, 2019. https://www.pewresearch.org/internet/2019/10/28/experts-optimistic-about-the-next-50-years-of-digital-life/ (accessed Jun. 25, 2020).

Supportive Material

B

This appendix contains materials that support the PhD thesis.

B.1 Publications and Keywords

This part contains the association between the publications included in this thesis and the keywords.

Publication 1
HealthMedia @ Multimedia 2019

Publication 2
MHC @ UbiComp 2018

Publication 3
LDC @ UbiComp 2019

Publication 4
MobiSPC 2020

Publication 5
DH 2018

Group 1
Challenges and Opportunities

Group 2
Designs Leveraging mQoL-Lab

Group 3
Designs of a Mobile App

Outcome

PRO

Behaviours

Nutrition
Physical Activity
Sleep

Health

Anxiety
Depression
Memory

Quality of Life

Health-Related QoL
Social Support

TechRO

Biomarkers

Heart Rate
Physical Activity
Sleep

Review

Participation

Challenges
Motivation Factors
Opportunities

Quality of Life

Energy
Fatigue

Model

Participation

Personalized Stories

Co-Calibration

coQoL

Operationalization

Setting

Community
In-Situ

Platform

mQoL-Lab

Artefact

Chatbot
Mobile App
Web App

Group 4
PRO Physical Activity

Group 5
PRO QoL and TechRO Heart Rate

Group 6
coQoL of numerous PRO-TechROs

Publication 6
JPM 2020

Publication 7
MobiSPC 2020

Publication 8
ISOQOL 2020

Publication 9
Quantifying QoL 2020

B.2 Publications and Scientific Contributions

This part contains the association between the publications included in this thesis and scientific contributions.

Publication 1
HealthMedia @ Multimedia 2019

Publication 2
MHC @ UbiComp 2018

Publication 3
LDC @ UbiComp 2019

Publication 4
MobiSPC 2020

Publication 5
DH 2018

Group 1
Motivation to participate review

Group 2
Mobile app design for studies

Scientific Contributions

Thoroughly explores a research area

Motivation for participation in human studies
Energy and fatigue Quality of Life facet

Provides empirical data

Physical and psychological PROs in seniors (raw)
Digital biomarker TechROs in seniors (aggregate)
Patterns of PRO-TechRO statistical correlations

Produces a computational model

coQoL co-calibration model for PROs and TechROs

Produces a presentational model

Personalized stories model to motivate participation

Provides a unifying framework

Mobile app design for longitudinal human studies
Chatbot design for longitudinal health studies

Designs and develops a tool

Module for questionnaire data collection PRO (CoME)
Module for wearable data collection TechRO (CoME)
mQoL-Lab platform for human studies (QoL Lab)

Group 3
PROs-TechROs co-calibrations

Publication 6
JPM 2020

Publication 7
MobiSPC 2020

Publication 8
ISOQOL 2020

Publication 9
Quantifying QoL 2020

B.3 Mobile Apps Comparison

This part contains a table that highlights similar paragraphs from the descriptions of the mQoL design [33], Apple app [31], and Google app [34].

Begin of Table B.1: Similarities between mobile apps: mQoL design, Apple Research, and Google Health Studies

Application	**mQoL**	**Apple Research app**	**Google Health Studies**
Status	Designed, August 2018	Released, October 2019	Released, December 2020
Motivation	"*mQoL is a mobile health platform [that facilitates] the deployment of much-needed longitudinal, multidimensional, evidence-based studies that provide high value in terms of the collected datasets. mQoL is designed to accommodate exploratory and interventional studies grounded in medical evidence.*" [33]	"*Your participation will enable innovative research that would have been all but impossible before. You can contribute to groundbreaking research studies. Simply download the Research app and enroll in a study. Humanity says thank you.*" [31]	"*Google Health Studies aims to create opportunities for more people to participate in health research. [It] lets you securely contribute to health research studies with leading institutions, right from your phone. By contributing, you'll represent your community and start improving the future of health for everyone.*" [34]
Principal entity: the study	"*The central concept of mQoL is the study, that acts as a research template. mQoL simplifies study deployment, by providing a platform that provides a format for designing studies, helping researchers worry less about app maintenance or study survival.*" [33]	"*We invite you to join one or more studies and make your mark on human health.*" [31]	"*Volunteer for multiple studies in one app.*" [34]
Patient-reported outcomes (PROs)	"*mQoL allows studies to design and schedule self-reported surveys and request access to [a set of shared] self-reported surveys. [Shared self-reported surveys include] a Quality of Life survey, a demographic survey, and a health survey.*" [33]	Separate research frameworks exist, e.g., ResearchKit [28] and HealthKit [29].	"*Self-report symptoms and other data. If you participate [in the first study], you'll provide data to help researchers understand how demographics, health history [and more] contribute [to the outcome of the study].*" [34]

Continuation of Table B.1: Similarities between mobile apps: mQoL design, Apple Research, and Google Health Studies

Application	mQoL	Apple Research	Google Health Studies
Technology-reported outcomes (TechRO)	"*mQoL uses device-reported individual health and smartphone usage data. Researchers can obtain rich behavioural datasets by retrieving longitudinal, multidimensional behavioral markers in time and context from participants.*" [33]	"*Devices like Apple Watch and iPhone capture meaningful health information, including signals from your heart, your level of motion and activity, and your sound exposure levels throughout the day.*" [31]	"*If you participate [in the first study], you'll provide data to help researchers understand how behavior and mobility patterns contribute [to the outcome of the study].*" [34]
Participant empowerment	"*Participants receive personalized, timely, and contextual information from studies helping them monitor, observe, and reflect upon daily life and its long-term health and QoL consequences.*" [33]	"*It will help Apple to create even more empowering technologies.*" [31]	"*Track your information with digital health reports. Learn research findings from the studies you participate in.*" [34]
Data storage	"*Health-related data is generated, measured, and collected on the device, continuously, unobtrusively, and independently of the app. For pseudonymous data retrieval, upon installing mQoL, the participant sets up a token. This token (and no other personal information) will identify the data retrieved from the participant.*" [33]	"*Any data collected through the Research app will be encrypted if you have a passcode set on your device. Apple will not have access to any contact information or other identifying data.*" [31]	"*Your personal information is kept on your device.*" [34]

Continuation of Table B.1: Similarities between mobile apps: mQoL design, Apple Research, and Google Health Studies

Application	mQoL	Apple Research	Google Health Studies
Data sharing	"*[Data] can be retrieved by each study upon consent. For retrieving data, each study requests only the most granular data types it needs, e.g., physical activity → walking → steps → daily → count. [Researchers can obtain] consented, pseudonymized, and structured datasets.*" [33]	"*What you choose to share for research, and with which study, is controlled entirely by you. Apple will not have access to any contact information or other identifying data that you provide through the Research app.*" [31]	"*Researchers only see aggregated study data combined from all participants. This allows researchers to collect the information needed to advance the study without seeing individual details.*" [34]
Personalization	"*When the number of studies increases in our app, we plan to design an onboarding feature to help participants choose those that suit their interests and can benefit them most.*" [33]	-	"*Volunteer for studies that matter to you and represent your community.*" [34]
Current studies	-	*Apple Heart and Movement Study* (Apple, American Heart Association, Brigham and Women's Hospital, Harvard Medical School), *Apple Women's Health Study* (Apple, Harvard T. H. Chan School of Public Health, National Institute of Environmental Health Sciences), *Apple Hearing Study* (Apple, University of Michigan) [31].	*Respiratory Health Study* (Boston Children's Hospital and Harvard Medical School) [34].

Tab. B.1.: Similarities between mobile apps: mQoL design, Apple Research, and Google Health Studies. Apple and Google descriptions selected and quoted *verbatim* from their respective websites. Bracket text added by the author of the thesis.

B.4 Academic Curriculum Vitae

This part contains the author's academic curriculum vitae.

Vlad Manea

Academic Curriculum Vitae

Contact Information

Quality of Life Technologies Lab, Section for Human-Centered Computing, Department of Computer Science, Faculty of Science, University of Copenhagen

Sigurdsgade 41, 2200 Copenhagen, Denmark

manea@di.ku.dk | personal website | google scholar | orcid

Education

PhD fellow in Computer Science at the University of Copenhagen
January 2018 – Present | Copenhagen, Denmark

Part of the Quality of Life technologies lab (Professor Katarzyna Wac), designing, developing, and evaluating emerging technologies that use data to benefit individuals' life quality. University of Copenhagen is No. 1 in the EU in computer science (Shanghai).

- Research in computer science, ubiquitous computing, and mobile health. Co-calibrated behaviours and health using wearables, surveys, and statistical data analysis. Published the results in a high-quality journal (impact factor 4.433, rank 10/102, quartile Q1). Initiated and co-organized the Longitudinal Data workshop at the top UbiComp conference.

- Research and development in the Quality of Life technologies lab. Co-supervised or advised master theses (5 graduate students, average grade 11.6/12). Implemented software used in research projects at the Universities of Geneva and Stanford. Reviewer or sub-reviewer for several mobile health-related conferences and journals.

- Administration in the Section of human-centred computing, Department of computer science. Served as the news communication liaison between the Section and the Department.

Completed six months of course in ubiquitous computing, statistics, machine learning, and related disciplines.

Master of Science in Computer Science at the University of Copenhagen
August 2013 – April 2016 | Copenhagen, Denmark

Focused on computer science innovation, machine learning, and algorithms. Participated by selection in the University of Copenhagen Innovation Hub. Defended a thesis in tree algorithms: Algorithmic Experiments for Online Matching.

Bachelor of Science in Computer Science at the University of Iasi
October 2008 – June 2011 | Iasi, Romania

Participated in the regional Intercollegiate Programming Contest (6 countries). Defended a thesis in information security and anonymity: DC and Mix Networks. Completed the bachelor program on the rank 1 / 196 graduates.

High School of Computer Science Iasi
September 2004 – June 2008 | Iasi, Romania

Real profile, mathematics and informatics, intensive informatics. Participated three times in the National Olympiad of Informatics, obtained a bronze medal, and qualified two times for the draw for international teams after the Olympiad. Obtained several national awards or mentions in educational software and database modelling.

Publications

Quality of Life

Co-calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method
Vlad Manea, Katarzyna Wac

Journal of Personalized Medicine, MDPI, 2020, Basel, Switzerland. Impact factor 4.433, rank 10/102 (Q1) in Health Care Sciences and Services. 41 pages | DOI | Slides

Energy and Fatigue. Classification and Assessment of Energy and Fatigue using Subjective, Objective, and Mixed Methods towards Health and Quality of Life
Natalie Solomon, Vlad Manea

Chapter in: Katarzyna Wac, Sharon Wulfovich (eds.), Quantifying Quality of Life: Incorporating Daily Life into Medicine, Springer, Cham, forthcoming. 30 pages.

Using Consumer-Friendly Wearables to Associate Patient-Reported Quality of Life and Tech-Reported Physical Activity and Sleep in Healthy Seniors
Vlad Manea, Katarzyna Wac

Poster Conference of the International Society for Quality of Life Research, ISOQOL 2020, online | DOI | Poster

Using Consumer-Friendly Wearables to Associate Patient- and Technology-Reported Physical Activity in Healthy Seniors
Vlad Manea, Allan Berrocal, Katarzyna Wac

International Conference on Mobile Systems and Pervasive Computing, MobiSPC 2020, Leuven, Belgium. 8 pages | DOI

mQoL Lab: Step-by-Step Creation of a Flexible Platform to Conduct Studies Using Interactive, Mobile, Wearable and Ubiquitous Devices
Allan Berrocal, Vlad Manea, Alexandre De Masi, Katarzyna Wac

International Conference on Mobile Systems and Pervasive Computing, MobiSPC 2020, Leuven, Belgium. 9 pages. **Nominated for the best project pitch award at the University Hospitals of Geneva Innovation Day** | DOI

Towards Personalizing Participation in Health Studies
Vlad Manea, Mads Schnoor Hansen, Ece Elbeyi, Katarzyna Wac

Workshop on Multimedia for Personal Health and Health Care, HealthMedia 2019. In conjunction with the ACM international conference on Multimedia, MM 2019, Nice, France. 8 pages | DOI | Slides

WellCo: Wellbeing and Health Virtual Coach
Vlad Manea, Katarzyna Wac

ERCIM News 118, Special Theme: Digital Health, July 2019. 2 pages | Edition

Call for Papers: Workshop on Longitudinal Data Collection in Human Subject Studies
Vlad Manea, Allan Berrocal, Alexandre De Masi, Naja Holten Møller, Katarzyna Wac, Hannah Bayer, Sune Lehmann, Euan Ashley

Call for Papers for the Workshop on Longitudinal Data Collection in Human Subject Studies, LDC 2019. In conjunction with the ACM international joint conference on Pervasive and Ubiquitous Computing, UBICOMP 2019, London, United Kingdom. 4 pages | DOI | Website | Twitter

mQoL: Mobile Quality of Life lab: from Behavior Change to Quality of Life
Vlad Manea, Katarzyna Wac

Workshop on Mobile Human Contributions, MHC 2018. In conjunction with the ACM international joint conference on Pervasive and Ubiquitous Computing, UBICOMP 2018, Singapore. 6 pages | DOI | Slides

Digital Health Tools for Chronic Illness and Dementia Risk Assessment in Older Adults
Sofia Laghouila, Vlad Manea, Vero Estrada-Galiñanes, Katarzyna Wac

Annual meeting and scientific sessions of the Society of Behavioral Medicine, Annals of Behavioral Medicine. SBM 2018, New Orleans, LA, United States | DOI

Assessing Chronic Illness Risk in Older Adults via Personal Digital Health Tools
Vlad Manea, Vero Estrada-Galiñanes, Katarzyna Wac

Poster ACM International Digital Health Conference. DH 2018, April 2018, Lyon, France | Poster

mQoL: Mobile Quality of Life lab
Vlad Manea, Vero Estrada-Galiñanes, Katarzyna Wac

Poster and demo ACM international Digital Health conference. DH 2018, April 2018, Lyon, France. **Nominated for the Innovation Prize in the category of the best data-driven innovation** | Poster | Slides | Demo

Healthcare Coordination

Balancing Priorities: A Field Study of Coordination in Distributed Elder Care
Troels Mønsted, Andreas Johansen, Frederik Lauridsen*, Vlad Manea*, Konstantin Slavin-Borovskij**

Hawaii international conference on system sciences. HICSS 2016, Kauai, HI, United States. 8 pages | DOI | Slides | * equal contribution

IT-Supported Coordination of Elder Care
Troels Mønsted, Andreas Johansen, Frederik Lauridsen*, Vlad Manea*, Konstantin Slavin-Borovskij**

Workshop on Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design. In conjunction with the European conference on computer-supported cooperative work, ECSCW 2015, Oslo, Norway. 5 pages | DOI | Article | * equal contribution

Improving Coordination of Care Centers for the Elderly through IT Support
Andreas Johansen, Frederik Lauridsen*, Vlad Manea*, Konstantin Slavin-Borovskij*, Troels Mønsted*

Poster ACM conference on computer-supported cooperative work. CSCW 2015, Vancouver, BC, Canada | DOI | Poster | * equal contribution

Software Engineering

Augmented Reality
Andrei Arusoaie, Ionuț Cristei, Cătălin Chircu, Andrei Livadariu, Vlad Manea, Adrian Iftene

International Symposium on numeric and symbolic algorithms for scientific computing. SYNASC 2010, Timisoara, Romania. 8 pages | DOI

Educational Software

.campion Educational Archive
Vlad Manea, Emanuela Cerchez, Marinel Șerban

National conference on virtual learning. CNIV 2009, Iasi, Romania. **Excellence award**. ISSN 1842-4708. 8 pages | Article | Award

Counselling Class Educational Portal
Vlad Manea, Alexandru Periețanu, Vlad Constantinescu, Claudia Cărăușu, Emanuela Cerchez

National conference on virtual learning. CNIV 2008, Constanta, Romania. **Intel Education award**. ISSN 1842-4708. 8 pages | Article | Award

Talks

Co-calibrating Physical and Psychological Outcomes and Consumer Wearable Activity Outcomes in Older Adults: An Evaluation of the coQoL Method
Vlad Manea, Katarzyna Wac

Meeting of the Health Informatics group, Sections of Human-Centered Computing and Software, People, Data, and Society, Department of Computer Science, Faculty of Science, University of Copenhagen. November 2020, Copenhagen, Denmark | Slides

Quality of Life at Your Fingertips
Vlad Manea, Jamshed Gill, Katarzyna Wac

Meeting of the Health Informatics group, Sections of Human-Centered Computing and Software, People, Data, and Society, Department of Computer Science, Faculty of Science, University of Copenhagen. October 2019, Copenhagen, Denmark | Slides

Quality of Life
Vlad Manea

Can Data Decide Your Health – Human / Legal Perspectives?
University of Copenhagen Innovation Hub. Copenhagen, Denmark, November 2018 | Event | Slides

Software

Quality of Life Wearable Data Collection System
2018 – 2020 | Copenhagen, Denmark

Wearable data collectors for the Quality of Life technologies lab. Implemented data collectors for most streams in the manufacturers' web APIs. Added schedulers, historic data collectors, visualization dashboards, data exporters, and more. Collects data from Fitbit and Withings.

Physical Health State Assessment Module
2018 – 2020 | Copenhagen, Denmark

Module integrated into the H2020 WellCo European project's application. Collects data from the WithingsAPI and stores it in the application's unified data store. Confronts guidelines with questionnaires and wearable data on physical activity and sleep. Identifies factors from the literature which are likely to impact the overall risk of disease. Released with extensive test and documentation, processing data continuously at HI-Iberia in Madrid.

Chronic Disease Risk Awareness Module
2018 – 2020 | Copenhagen, Denmark

Module integrated into the H2020 WellCo European project's application. Quantifies the direct risk factors for cardiovascular, pancreatic, and pulmonary disease. Generates "if you continue like this..." and "what if?" alternative behavioural risk scenarios. Provides the minimal behaviour changes needed to reduce the modifiable risk in each situation. Released with an extensive test (coverage 94%) and documentation, processing data at HI-Iberia in Madrid.

Physical Health State Assessment Module
2018 – 2019 | Copenhagen, Denmark

Module integrated into the AAL CoME European project's application. Implemented the data models and out-come scoring of 8 clinical instruments. Quantified risk factors: anxiety, depression, health-related life quality, memory, nutrition, physical activity, sleep, and social support. Served the questionnaires and risks as an API consumed by a web app running in Madrid. Released with extensive test and documentation, collected and exposed data at the University of Geneva.

Review

Review for the Conference on Pervasive Technologies in Healthcare PervasiveHealth
January 2020 | Copenhagen, Denmark

Review for the Journal of Medical Internet Research mHealth uHealth
October 2019 | Copenhagen, Denmark

Review for the Conference on Pervasive Technologies in Healthcare PervasiveHealth
January 2019 | Copenhagen, Denmark

Review for the Conference on Pervasive Technologies in Healthcare PervasiveHealth
February 2018 | Copenhagen, Denmark

Supervision

Co-supervisor for a master thesis on Engaging Participants in the Recruitment Phase of Human Subject Health Studies – mQoL-chat: a Chatbot Approach
Spring 2019 | Copenhagen, Denmark

Mads Schnoor Hansen

Master thesis in Computer Science. 2019. Department of Computer Science, Faculty of Science, University of Copenhagen, Denmark.

Advisor for a master thesis on Designing for Participation in Longitudinal Health and Well-being Studies
Spring 2020 | Copenhagen, Denmark

Alba Kejser Perez, Cecilie Rosentoft, Elisabeth Brinth Refstrup

Master thesis in Communication and IT. 2020. Department of Communication, University of Copenhagen, Denmark.

Advisor for a master thesis on Framework for Modelling Consistency in Behavioural Data Collected with Wearables
Spring 2020 | Copenhagen, Denmark

Kirke Kjellberg

Master thesis in Computer Science. 2020. Department of Computer Science, Faculty of Science, University of Copenhagen, Denmark.

Teaching

Assistant Lecturer the University of Copenhagen
Fall 2016 | Copenhagen, Denmark

Graded work at the Advanced Computer Systems course for the M.Sc. in computer science at the Department of Computer Science, University of Copenhagen, 100 students.

Invited Assistant Lecturer at the University of Copenhagen Innovation Hub
Winter 2015 | Copenhagen, Denmark

Invited lectures on Prototyping for the Knowledge-Based Entrepreneurship course for the M.Sc. in computer science at the Department of Computer Science, University of Copenhagen and the University of Copenhagen Innovation Hub, 15 students | Slides

Teaching Assistant at the IT University of Copenhagen
Fall 2013 | Copenhagen, Denmark

Held seminars and graded work at the Introductory Java Programming course for the M.Sc. in computer science at the Department of Systems and Software, the IT University of Copenhagen, 45 students.

Teaching Assistant at the University of Iasi
Fall 2011 | Iasi, Romania

Held seminars and graded work at the Algorithms and Programming course for the B.Sc. in computer science at the Faculty of Informatics, the University of Iasi, 30+30 students in my group, 400+ overall.

University Student Lecturer at the University of Iasi
October 2009 – December 2011 | Iasi, Romania

Held seminars at the Algorithms course for high school students preparing for the Olympiad in Informatics at the City Center of Excellence Iasi. 15 students every year.

Industry Experience

Head of Data Analysis at Steps
January 2018 – October 2018 | Copenhagen, Denmark | Part-Time

Steps made mental health fun, actionable, and accessible, by providing users with insights and tools to build social confidence one step at a time - while helping their institutions support them in their journey.

Part of the technical team of the startup and leading the data analysis track. Collected, analyzed, and reported usage data from two customer pilot deployments. Built and refactored flows, and introduced design patterns to reduce work in our chatbot. Participated by selection in the University of Copenhagen Innovation hub.

Software Developer at Alipes Capital
August 2016 – December 2017 | Copenhagen, Denmark | Full-Time

Alipes Capital is an innovative company working in the financial technology sector. It uses market news releases to trade fast: "every millisecond counts for victory!"

Part of the development team, implementing internal tools for high-frequency trading. Developed a link predictor running in 1 second, discovering 100 valid links per quarter. Optimized our internal workbench, decreasing the runtime from 40 to 5 minutes. Implemented custom file parsers based on heuristics, which ran in under one millisecond. Designed and implemented algorithms, data structures, and heuristics on a daily basis.

Co-Founder and Head of Engineering at Habitlab
October 2015 – July 2016 | Copenhagen, Denmark | Full-Time

Habitlab built a habit management platform aimed at helping individuals, organizations, and the healthcare sector use good habits as a catalyst for change.

Led the technical tracks in the startup: products, people, and processes. Built and released a software ecosystem: one framework, two mobile apps, and one web app. Achieved a fast feature release cadence of less than 1 week, close to Apple's review time. Led 5 part-time developers, set up and followed up on personalized development plans. Managed agile processes: sprints, issue tracking, pair programming, code review, versioning. Built two mobile apps (Habitlab and Physio). Qualified for the Thinkubator incubator (7% acceptance rate).

Student Developer at e-conomic
December 2013 – September 2015 | Copenhagen, Denmark | Part-Time

e-conomic is the most intuitive accounting program in Denmark.

Part of the machine learning, platform, and user interface teams during my master. Implemented cloud prototypes in Azure machine learning to classify account types. Developed microservices that broadcasted events between analytics platforms. Upgraded libraries, decreasing the build time by 35% on developer machines. Wrote an

interpreter that converted 1.200 tests between javascript dialects. Worked within an agile process using scrum and kanban.

Software Engineering Intern at Google
June 2013 – September 2013 | Mountain View, CA, United States | Full-Time Internship

Part of the Google+ external APIs team. Developed the API to create Google+ accounts used by estimated 100.000 Android devices. Organized a Q&A event about Google that gathered an engineering director and 30 interns.

Software Development Engineer at Microsoft
November 2012 – May 2013 | Copenhagen, Denmark | Full-Time

Part of two teams for the Microsoft Dynamics CRM Marketing Pilot 15 product: user experience refresh and localization. Built user interface controls recurring on all pages or on the main page. Refreshed 20+ existing user interface controls to the Microsoft Modern Design. Gently guided our user experience designer to deliver on time and with quality. Reconciled date and time representations across the product stack and features. Worked within an agile process using scrum, kanban, and pair coding (both lead and led).

Software Development Engineer in Test at Microsoft
January 2012 – October 2012 | Copenhagen, Denmark | Full-Time

Part of two teams for the Microsoft Dynamics NAV 2013 product: client core and application. Created environments in virtual machines to conduct manual test scenarios. Developed automated tests as part of team acceptance test-driven feature development. Worked within an agile process using scrum and kanban.

Software Development Engineer Intern at Amazon
December 2010 – May 2011 | Iasi, Romania | Part-Time Internship

Part of the platform team during my bachelor studies. Implemented scripts to persist internal frequent query results. Implemented a cloud app for ordering food.