CALL H2020-SC1-FA-DTS-2018-2020
Trusted digital solutions and Cybersecurity in Health and Care
TOPIC DT-TDS-01-2019
Smart and healthy living at home

SMART BEAR

"Smart Big Data Platform to Offer Evidence-based Personalised Support for Healthy and Independent Living at Home"

D53 – Evaluation Framework

Due date of deliverable: 31/08/2022
Actual submission date: 31/08/2022

Grant agreement number: 857172
Start date of project: 01/09/2019

Lead contractor: CNR
Duration: 60 months

Project funded by the European Commission within the EU Framework Programme for Research and Innovation HORIZON 2020	
Dissemination Level	
PU = Public, fully open, e.g. web	✓
CO = Confidential, restricted under conditions set out in Model Grant Agreement	
CI = Classified, information as referred to in Commission Decision 2001/844/EC.	
Int = Internal Working Document	

D12.2 (D53) – Evaluation Framework

Editors
Eleftheria Iliadou [NKUA]
Athanasios Bibas [NKUA]
Luigi Gallo [CNR]
Ioannis Basdekis [STS]

Contributors
Agnese Augello [CNR]
Ioannis Basdekis [STS]
Giuseppe Caggianese [CNR]
Luigi Gallo [CNR]
Yao Chen [CHU Lille]
Eleftheria Iliadou [NKUA]
Ioannis Leontsinis [NKUA]
Mircea Marzan [ANA]
Christos Nikitas [NKUA]
Andreia Pimenta [SRS]
Stefano Silvestri [CNR]
Luiza Spiru [ANA]
Andreea Stamate [ANA]

Reviewers
Stela Musteata [FCSR]

The work described in this document has been conducted within the project SMART BEAR, started in September 2019. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857172

©Copyright by the SMART BEAR Consortium.

Executive Summary

This deliverable aims to present the SMART BEAR evaluation framework, and thus the refined set of key performance indicators (KPIs) and the criteria and methodology set to evaluate the SMART BEAR platform. In the following chapters, we present:

1. the refined set of clinical and technical KPIs, along with their selected metrics,
2. the implementation plan of the MAFEIP framework in SMART BEAR, and
3. the design and methods of the SMART BEAR usability evaluation framework.

Once the large-scale study is launched, the results of this evaluation framework will be provided in the cross-pilot deliverables.

Contents

1 Introduction **7**
2 Clinical KPIs **10**
2.1 General 10
2.2 Hearing loss 11
2.3 CVDs 11
2.4 Frailty 12
2.5 Mild Cognitive Disorder 13
2.6 Balance Disorders 14
2.7 Low Back Pain 15
3 Technical iKPIs and KPIs **18**
3.1 Technical Milestones 22
3.2 Technical KPIs: Progress 23
4 MAFEIP Evaluation framework **28**
4.1 Models and monitored parameters 29
4.2 General data 33
4.3 Model Input data 36
4.3.1 Set-up 36
4.3.2 Probabilities 36
4.3.3 Cost inputs 37
4.4 Simulation 37
5 Usability Evaluation Framework **38**
5.1 Methodology 38
5.2 Stage 1 Procedure 39
5.3 Stage 2 Procedure 40
5.4 Ethical Guidelines and Requirements 41
6 References **42**
7 Appendix **48**
7.1 Usability Evaluation Framework 48
7.1.1 Summary of the evaluation tools retrieved in the literature 48
7.1.2 Stage 1 50
7.1.2.1 Descriptions of tasks for the app evaluation 50
7.1.2.2 Descriptions of tasks for the dashboard evaluation 51
7.1.2.3 Descriptions of the evaluation tools 53
7.1.3 Stage 2 56
7.1.3.1 Descriptions of the evaluation tools 56

List of tables

Table 1. Summary of iKPIs 7
Table 2. iKPIs concerning low back pain in the Pilot of Pilots 16
Table 3. Objectives and present status 18
Table 4. Milestones and present status 22
Table 5. Summary of KPIs and current status 23
Table 6. Mild Cognitive Disorders parameters for MAFEIP models 30
Table 7. First group of CVD parameters for MAFEIP models (Hypertension, Coronary Artery Disease, Arrhythmias) 30
Table 8. Second group of CVD parameters for MAFEIP models (Hypertension, Coronary Artery Disease, Arrhythmias) 31
Table 9. Second group of CVD parameters for MAFEIP models (Heart Failure) 31
Table 10. Frailty parameters for MAFEIP model 32
Table 11. Mild Depression parameters for MAFEIP model 32
Table 12. Low back pain parameters for MAFEIP model 32
Table 13. Balance Disorders parameters for MAFEIP model 33
Table 14. Hearing Loss parameters for MAFEIP model 33
Table 15. Tools retrieved in the literature. 48
Table 16. Task 1@App: Diet App - Monthly Habits Questionnaire 50
Table 17. Task 2@App: Diet App - Daily food intake 50
Table 18. Task 3@App: OTAGO/rehabilitation - Monthly Questionnaire Scenario 51
Table 19. Task 4@App: Medication Intake 51
Table 20. Task 1@Dashboard: Setting of a patient's health habits 52
Table 21. Task 2@Dashboard: Medications Setting 52
Table 22. Task 3@Dashboard: Questionnaire Completion 53
Table 23. Task 4@Dashboard: Notebook (AI) 53
Table 24. After Scenario Questionnaire 54
Table 25. Post Study System Usability Questionnaire. * not to show to the subjects 54
Table 26. Usability Metric for User Experience Questionnaire 56
Table 27. mHealth App Usability Questionnaire for Clinician. * not to show to the subjects 56
Table 28. mHealth App Usability Questionnaire for Participants: * not to show to the subjects 57
Table 29. Net Promoter Scale 58
Table 30. UTAUT Items 59

List of figures

Figure 1. MAFEIP Markov Models 29
Figure 2. Example of "Setting and target population" data 34
Figure 3. Example of "About your intervention" subsection 35
Figure 4. Example of "Evidence" subsection 35
Figure 5. Participant Involvement in the two Evaluation Stages 39
Figure 6. The two Evaluation Stages: a redesign can be carried out according to the obtained results 39
Figure 7. UTAUT model (Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly) 59

1 Introduction

SMART BEAR aims to provide its participants with a state-of-the-art technological solution that will enable them to deal better with their comorbidities. SMART BEAR aims at improving specific domains of participants' living and achieving an important impact. These aims reflect the following list of KPIs (Table 1 and Table 2). The rationale behind each KPI is provided later in the manuscript.

Table 1. Summary of iKPIs

iKPI ID	DESCRIPTION AND TARGET VALUE	VALIDATION
iKPI 1	Number of healthcare standards contributed by *SMART BEAR*. Targeted value: At least 2 standards.	Reported in D13.5.
iKPI 2	Number of external partnerships promoting the deployment of the final product in related markets. Targeted value: At least 2.	Identified synergies and cross-collaborative side-projects. Reported in D13.5, D13.6.
iKPI 3	Number of meetings with relevant stakeholders. Targeted value: At least 6 per year.	Reported in D13.6 .
iKPI 4	Decrease of total individualized cardiovascular risk score for users with non-established CVD as defined by the European Society of Cardiology Targeted value:<5%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 5	Increase of prospective potential of evidence-based platform-agnostic technology in the healthcare industry. Targeted value: >10%.	Impact assessment and outcomes of pilots-based validation as reported in D12.5, D13.5.
iKPI 6	Improved percentage of adherence in prescribed medical treatment. Targeted value: Increase by 20%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 7	Decrease in patient-to-doctor / patient to Audiologist visits. Targeted value: 20% fewer visits annually.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 8	Decrease in disease related number of admittances and hospitalisation time for participants suffering from ischaemic heart disease and heart failure. Targeted value: 15%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 9	Achievement and maintenance of hypertension treatment targets as defined by the European Society of Hypertension.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 10	Achievement of increased and optimised hearing aid use through objective recordings and consequent analytics of the SMART BEAR platform. Targeted value: 25%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 11	Improve users' social well-being. Targeted value: >20%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 12	Participants will meet the following WHO guideline: Older adults should do at least 150 minutes of moderate-intensity aerobic physical activity throughout the week or do at least 75 minutes of vigorous intensity aerobic physical activity throughout the week or an equivalent combination of moderate- and vigorous-intensity activity.	Outcomes of pilots-based validation reported in D8.4, D9.4, D10.4, D11.1, D12.5.

iKPI 13	Increase in EQ-5D index.* Targeted value: >0.5 SD (standard deviation) change EQ-5D replaced original quality of life questionnaire QALY, in agreement with current literature and to enable the implementation of MAFEIP tool.	Outcomes of final cross-pilot evaluation reported in D12.5
iKPI 14	Increase in user interaction satisfaction index (QUIS). Targeted value: >75	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 15	Number of presentations at scientific events. Targeted value: >20.	Reported in D13.5.
iKPI 16	Number of meetings with research institutions. Targeted value: >20.	Reported in D13.5, D13.6 .
iKPI 17	Increase the degree of usability of the platform. Targeted value: UMUX average score > 80 (UMUX score ranges between 0 and 100)* * This new IKPI replaced the previous: Increase the degree of usability of the platform. Targeted value: >80 in SUS and >80% in TAM . The changes are motivated in Section 5.1.1.	Usability Metric for User Experience (UMUX).
iKPI 18	Faster and smarter identification of exploitable errors and data breaches in a healthcare platform. Targeted value: Number of identified errors and breaches <10%.	Functional, Integrations and Penetration testing Reports, Integrity, Confidentiality and Availability, Security Assurance Reports (reported in D5.7) .
iKPI 19	Increase of HA usage >10% than general population.	Outcomes of final cross-pilot evaluation reported in D12.5
iKPI 20	Increase of GHABP score at the end of the project.	Outcomes of final cross-pilot evaluation reported in D12.5 .
iKPI 21	The monitoring solution will help keep drop-out rates due to various reasons (interoperability, usability, data security/privacy issues, user acceptance) low (<20%).	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 22	By the end of the first 6-month of the pilot participants who comply with the suggested cognitive training program will show an increase in MoCA score by at least 1 point.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 23	By the end of the first 6-month of the pilot, participants who comply with the suggested cognitive training program and physical activity intervention will show an increase in GDS score by at least 10%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 24	By the end of the first 6-month of the pilot, participants who comply with the suggested cognitive training program will show an increase in iADL score by 1 point.	Outcomes of final cross-pilot evaluation reported in D12.5.

iKPI 25	By the end of the first 6-month of the pilot, participants who comply with the suggested intervention program will show a decrease in EFS score by 1 point.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 26	Decrease the number of falls and injuries related to falls. Targeted value: 25%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 27	Decrease the fear of falling indicated on Falls Efficacy Scale-International (FES-I) questionnaire and increase balance confidence as indicated by Activities-specific Balance Confidence Scale (ABC) after the completion of the rehabilitation program by the end of rehabilitation program. Targeted value: 10%.	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 28	Increase postural stability indicated by Balance tests [Functional Gait Assessment (FGA) and mini Balance Evaluation Systems Test (miniBESTest)]. Targeted value : 10%	Outcomes of final cross-pilot evaluation reported in D12.5.
iKPI 29	Decrease drop-out rate after the completion of the rehabilitation protocol. Targeted value: 30%.	Outcomes of final cross-pilot evaluation reported in D12.5.

2 Clinical KPIs

2.1 General

Even though we have initially indicated a dropout rate of less than 5%, we have strong reasons to expect higher dropout rates. Recent evidence shows that the dropout rate in older population depends a lot on the health condition/s (Beller et al., 2022). For example, in mental health disorders, among older adults, the dropout rate is higher (i.e., depression 75% after 6 months due to "lack of relief/partial relief/worsening of illness) (Grover et al., 2022). Moreover, today we know that there are different aspects of health that predict dropout in a contrasting manner: better physical and cognitive functioning predicted decreased odds of dropout in women, whereas fewer chronic conditions, and thus a better health status in terms of less chronic conditions, predicted increased chances of study dropout. Thus, participants who have chronic conditions but are not negatively impacted by them are most likely to participate in future survey waves. Having chronic conditions, on the other hand, was associated with decreased odds of dropout. Having chronic conditions might increase one's need to talk about these issues and might thus increase one's willingness to participate in studies that address these topics (Beller et al., 2022). Considering that a high number of participants with different backgrounds, from several countries, will be involved in various factors are likely to impact and lead to a higher dropout rate. Furthermore, the majority of the studies in the literature report dropout rates of 11% and above over a 1-year framework (Viken et al. 2019). A recent meta-analysis by Meyerowitz-Katz et al. (2020) on 17 studies showed that the pooled dropout rate was 43% (95% CI 29-57). Out of the 17 studies 9 (53%) were randomized controlled trials, and 8 (47%) were observational trials; both types of studies covered a range of chronic diseases, as does the current study. Observational studies had a higher dropout rate of 49% (95% CI 27-70) than randomized control trials which had a 40% dropout rate (95% CI 16-63). The high variability between the studies is represented statistically in the high degree of heterogeneity (I2>99%). Various factors have been identified to be related to attrition from app-based interventions: behavioural characteristics, health status, literacy, age (younger participants drop out less) and level of education (postgraduates drop out less), and number of studies/interventions they are involved in (low attrition for those engaged in multiple interventions) (Meyerowitz-Katz et al., 2020, Grover et al., 2022). The results from these studies also suggest ways to improve engagement and attrition by using tailored messages to maintain high the users' interest and compliance. This is one of the main features of the SMART BEAR App, the users will receive personalized messages designed to aid with the management of their medical conditions. Additionally, we will implement support call centres to aid the participants during the study with any technical (or other) issues related to the use of the system. Taking into consideration the above-mentioned literature, but also the previously mentioned "improvement" that we aim to include in the study methodology, we expect a dropout rate below 20% over 1 year study (IKPI 22).

We are also targeting to improve the quality of life and social wellbeing of those participants that do not dropout by addressing their comorbidities with the use of one single technological solution and

taking into account physiological, medical, behavioural and environmental factors (Bellandi et al., 2022; Da Silva et al., 2021) (IKPI 11, 13 and 23).

2.2 Hearing loss

Debilitating hearing loss (HL) affects approximately 6% of the human population and by 2050 nearly 2.5 billion people are projected to have some degree of hearing loss (Śliwińska-Kowalska & Zaborowski, 2017). Only one out of five people in need of a hearing aid will eventually acquire one. The number of people that are efficiently using their hearing aids is even lower. Understanding the personal, behavioural, environmental, or other factors that correlate with the optimal conditions for hearing aid experience is a key point in improving patients' satisfaction and quality of life (Jilla et al, 2020). Duration of daily hearing aid use has been correlated with higher satisfaction and hearing aid benefit.

In SMART BEAR, during a 12-month period, we are making use of the capacity of modern hearing aids to provide dynamic logging of their operation and we combine this information with a plethora of medical, environmental, and social information for each user. This fact, along with the fact that SMART BEAR participants will have access to face-to-face and remote fine tuning and counselling sessions, according to their individualized needs lead us to target to increased hearing aid use adherence (IKPI-10, Targeted value: 25% less drop-outs than in general population), decrease of the dependence on Audiologist visits and repeated counselling (IKPI-7, Targeted value: 20% fewer visits than general population which is ~4 during the first 6 months, (Tecca, 2018)), increased hours of hearing aid usage (IKPI-19, Targeted value >10% than general population), and increased benefit from the hearing aid at the end of the project (IKPI-20, Target: Increase of GHABP score at the end of the project (Whitmer et al, 2014)).

2.3 CVDs

Cardiovascular disease represents one of the leading causes of morbidity, mortality, and disability globally. It affects all age groups but its dominance among the elderly people is profound. Societies in the contemporary era confront numerous challenging structural and functional changes. Population ageing constitutes a major issue on this aspect. In Europe, current predictions for the near future support a forthcoming increase of the elderly population as the demographic growth is slowing down while ageing accelerates. The constantly increasing morbidity rates in conjunction with the gradual lack of independence among the older citizens summarize the major burdens that the health and insurance systems are called to deal with. Towards the same direction, family pattern transmission to a looser structure amplifies the problem. This is easily understood as the less support the family can offer towards an elderly and partially dependent person, the more vulnerable the person becomes. A continuously growing vicious circle.

The CVDs section of the SMART BEAR projects aims to offer a smart, guideline directed aid to this vulnerable part of the population who often face multiple comorbidities, with personalized advice and or interventions aiming to promote active living, healthy nutritional and social habits. This intervention is consequently expected to enhance the safe and independent living and improve the quality of life.

One of the major goals of the project is to improve patients' behavior towards medical advice. This refers among others to adequate blood pressure control, dietary restrictions, adherence to medications, maintenance of adequate physical activity, abstinence from smoke products. It is well studied that treatment non adherence is one of the major reasons for pseudoresistant hypertension (Mancia et al., 2018), recurrent hospitalizations for heart failure decompensation, and worse cardiovascular outcomes. At the same time, there are several reports which endorse the usage of aids to improve adherence to medications and lifestyle modifications in a manner to improve overall quality of life. (Sheppard et al, 2021; Senoo et al, 2022; Tam et al, 2022; Tromp et al, 2022; Lukitasari et al, 2021; Wan et al, 2022; Ma et al, 2022; Zhang et al, 2019; Bhandari et al, 2022; Nguyen-Huynh et al, 2022; Ernawati et al, 2022). SMART BEAR offers the pioneer opportunity for all users to monitor their blood pressure levels on a daily basis for 12 months. Regarding optimal blood pressure levels, the SMART BEAR platform has adopted the ESH/ESC guidelines (Mancia et al, n.d.) in a way that the end-users will be able to get notifications regarding their Blood Pressure (BP) status and modify their lifestyle and or medication regime if their medical practitioner has advised them likewise in a way that participants will be helped to improve dietary habits and sustain optimal Body Mass Index (BMI).

The SMART BEAR project aims to improve the percentage of adherence in prescribed medical treatment by 20%, by means of personalized reminders via a prespecified application. The result will be audited in a monthly basis via the application and its effectiveness regarding the BP control apart from the out of office BP measurements taken at the patient's convenience, will be cross-checked periodically (6 months) over the prespecified medical review. The optimal BP control will result in a decrease of the number of non-scheduled visits due to hypertensive urgencies or emergencies (Angeli et al, 2020; Astarita et al, 2020; Leontsinis et al, 2021), especially in the population with ischemic heart disease. This effect will be measured by comparing the baseline self-reported values during history taking. In the same line, the project aims to reduce the number of non-scheduled visits due to volume overload in subjects with heart failure, which is a major issue in heart failure care as most of the patients encounter at least one urgent hospitalization per year accompanied by a significant deterioration in prognosis (IKPI 4,6,8,9, 12 and13).

2.4 Frailty

A review of the literature on interventions to prevent or reduce the level of frailty in older adults (Puts et al., 2017) included fourteen studies of which: 12 randomised controlled trials and 2 cohort studies (mean number of participants 260, range 51–610). The interventions included physical activity; physical activity combined with nutrition; physical activity plus nutrition and memory training; home modifications; rehabilitation (physical therapy plus exercise plus home modifications) and comprehensive geriatric assessment (CGA). The review showed that the interventions, that significantly reduced the number of frailty markers or the prevalence of frailty, included the physical activity interventions (all types and combinations), and rehabilitation. The CGA studies had mixed findings.

Only one of the 12 studies included in Puts et al. (2017) review used Edmonton Frail Scale (EFS) scores (which we will use in our study), which reduced by 1 point after 6 months intervention. We therefore also propose a reduction by 1 point on the EFS as an outcome of the SMART BEAR intervention (IKPI 25).

Stookey and Katzel (2020) reviewed the most recently published home-based exercise interventions for frail older adults living at home (8 studies). Research interventions consisted of a variety of modes (e.g., strength, strength/nutrition, strength/flexibility/balance/endurance), durations (e.g., 12 weeks

to 6 months), frequencies (e.g., 2–7 days/week), and delivery methods (e.g., volunteer-led, videos on a tablet, manuals/brochures). Investigators examined the effects of home-based exercise on a variety of outcomes to include feasibility, frailty status, physical performance, lean body mass, skeletal muscle mass, other physiological outcomes, mental health, nutritional status, and incidence of falls in frail. The results of this review demonstrate the feasibility and effectiveness of home-based exercise interventions to improve frailty, functional performance, nutritional status, and incidence of falls in frail older adults living at home.

2.5 Mild Cognitive Disorder

The existing reviews and previous meta-analyses on the effects of cognitive interventions in older adults have reported varying findings concerning the benefits of cognitive training. While several studies reported the benefits of cognitive training others have found little or no advantage (Peng et al., 2019).

In subjects with Mild Cognitive Impairments (MCI), a less recent review by Huckans et al. (2013) concluded that, according to the results of randomized controlled trials, efficacy of cognitive rehabilitation therapies for MCI can be categorized as encouraging but inconclusive.

A recent review, on 26 primary studies (Da Silva et al., 2021), 18 of these reported beneficial effects of educational and cognitive interventions for QoL, psychological well-being, and depressive symptoms in older adults without dementia and/or with MCI.

A recent Meta-Analysis on the effects of cognitive interventions in older adults by Yun et al., (2022) reviewing 38 studies with a total number of 5,238 subjects in experiment groups and 5,079 in the control groups, found that several cognitive-based interventions have been shown to be effective in improving cognitive function in older adults. Cognitive stimulation and cognitive training showed similar effects in the elderly with healthy cognition and those with MCI, whereas cognitive stimulation showed the greatest effect in the elderly with dementia. Total effect size of cognitive-based intervention on cognitive function of older adults was small to moderate (i.e., the maximum increase in MoCA score was 2 points). According to the study, when designing cognitive-based intervention it seems necessary to consider one session to be 60 minutes, scheduled twice per week, over a total operation period of 16–17 weeks.

A less recent Meta-Analysis on the generalizability of cognitive intervention effects to objective cognitive measurements, included 20 studies and found that Cognitive Computerized Training (CCT) showed low generalization of the trained aspects when assessed using cognitive testing (Simon et al., 2012).

Hill and colleagues (2016) reviewed several studies (e.g., 17 trials) on CCT applied to MCI population and concluded that CCT is a viable intervention for enhancing cognition (i.e., the overall effect on cognition in MCI across 17 trials was moderate: Hedges' g=0.35, 95% CI=0.20-0.51). From this review, it was noticed that participants of CCT groups improved significantly over the intervention period, while controls did not show any cognitive change. Most of the trials (70%) used an active control condition and the effects across active- and passive-controlled trials were comparable. The results of the MCI analysis are therefore robust and indicate a beneficial therapeutic role for CCT in this population.

Another recent meta-analysis by Hu and colleagues (2019), that included 12 studies on older adults with Subjective Cognitive Complaint and MCI, showed that Cognitive Computerized Training was most effective in cognitive rehabilitation, particularly in the subdomain of memory. Specific CCT was more

promising than non-specific CCT (i.e., computerized cognitive activities that do not have a treatment-oriented purpose) in patients with MCI.

A review (Gates et al, 2019), examining the effects of CCT compared to active or inactive controls on cognitive function in older adults with MCI, has produced contradictory results. Eight randomized controlled trials (total of 660 participants) were reviewed and CCT appeared to improve performance on global cognition, episodic memory and working memory compared to active controls. However, these results are based on very low-quality evidence.

In summary, the main benefits of CCT were improvements in memory, working memory, processing speed and attention. Indeed, CCT was found to be as effective as the traditional cognitive training, but less labor-intensive alternative, safe and efficacious. Furthermore, CCT had increased compliance in older adults, possibly because it is easy to access, can be used directly from home, is non-invasive, relatively inexpensive and does not require particular technological skills.

A systematic review of the literature on 13 studies with 7130 participants in total (Fan et al., 2019) examining whether cognitive training, independent of other interventions, can improve instrumental activities of daily living (iADL) performance in older adults with Mild Neurocognitive Disorder (MND) concluded that cognitive training may benefit in improving iADL performance in older adults with MND, but has no major effect on the independent living. Out of 13 studies reviewed in the metanalysis:

- Six reported a significant change on validated iADL assessment (i.e., five studies included older adults with normal cognition and one included subjects with MCI - the effect size was small = 0.15).
- Eleven studies showed improvement in measures of cognition.
- None of the studies described changes in independent living.

In conclusion, the reviews and the meta-analyses reported above, suggest significant positive effects of CCT on cognition and performance in instrumental activities of daily living in older adults with MCI. Therefore, CCT will be one of the SMART BEAR interventions provided to older adults with Mild Cognitive Disorder (IKPI 23-24).

2.6 Balance Disorders

Balance disorders are considered an epidemic in the older population with multiple negative consequences such as falls (i.e., the main consequence). Falls are considered the second leading cause of accidental deaths in the targeted population (CDC,2019); one out of three older adults falls annually. Vestibular deficits, cognitive impairment, sarcopenia, and vision impairment increase dramatically the risk of falling (Ambrose et al., 2013). A multi-factorial rehabilitation program could effectively decrease the fear of falling and the number of falls and injuries (Gillespie et al., 2012) such as increase postural stability (Skelton et al., 2005). Furthermore, we now know supervision increases adherence (Hall et al., 2022).

SMART BEAR participants with a balance disorder will be eligible for one of the two intervention options provided by the platform; a widely used well-documented protocol for falls prevention (OTAGO exercise program), provided by a fully automated and rule-based application or a beyond the state-of-the-art rehabilitation platform, provided via an augmented reality environment (HOLOBalance) for at least eight consecutive weeks. Decreasing physical inactivity and sedentary behaviour by engaging in a rehabilitation protocol targeting in motivation raising (via individualization of clinical decision making and virtual coaching) and combining with several bio-psycho-social

features of each participant will improve adherence (targeted value: 30% less drop-out rate than in general population), decrease in number of falls and injuries related to falls (Targeted value: 25% less falls and injuries related to falls than in general population), decrease fear of falling (targeted value:12% based on the FES-I validated questionnaire) and consequently increase of balance confidence (Targeted value:10% based on the ABC questionnaire) and increase postural stability (Targeted value: 10% on scores of FGA and miniBESTest (IKPIs 26, 27, 28 and 29).

2.7 Low Back Pain

Back pain is one of the leading causes of disability and reduced performance worldwide with an incidence prevalence ranging from 60 to 70%, (Airaksinen et al, 2006a), and it is being addressed in SMART BEAR Pilot of Pilots in the context of the synergy with SMART4HEALTH and HOLOBALANCE. Acute, subacute, and chronic back pain subtypes can be identified, although very often a clear diagnosis cannot be established, yet in most cases back pain is considered as a set of symptoms and often characterized as non-specific low back pain (Airaksinen et al, 2006b). Low back pain (LBP) treatment is a standardized procedure with clinical guidelines proposing specific interventions (e.g., physical, educational cognitive-behavioral, pharmacological and/or invasive procedures) with the main scope of achieving a multifactorial approach (Schiltenwolf et al, 2014; Schiltenwolf et al, 2017) and accomplishing long term behavioral changes.

Aspects of cognitive function, such as information processing and attention (Van Daele et al 2010; Boisgontier et al, 2013), seems to be affected in people with non-specific LBP. A recent systematic review (Kohl & Hansel, 2019) showed that people with non-specific LBP present postural control changes in combination with neuromuscular and biomechanical parameters, while especially for the elderly with chronic LBP, the impaired postural adjustments are leading scientists to recommend postural stability exercises as part of a rehabilitation program (Garcez et al, 2021).

A greater need for mental focus by people with non-specific LBP while performing dual tasks (Garcez et al, 2021) are increasing the events of falls and consequently the risk of injur (Shanbehzadeh et al, 2018). These findings could partially explain the increased risk of falling in people with back pain history (Marchall et al, 2016). After all, falls are one of the most common causes of fatal injury and LBP can potentially be added to the numerous risk factors already identified (e.g., vestibular disorders, cognitive impairment, physical inactivity) for increasing the incidence of falls (Pothula et al, 2004). An augmented reality platform (i.e., Holobalance), through a multi-sensory balance rehabilitation intervention (Gatsios et al, 2019), is being used for balance rehabilitation and fall prevention, in SMART BEAR project. Holobalance promotes falls prevention by improving the balance, as well as enhancing compliance and empowerment mobility behavioral changes in the long term.

Additionally, exercise can improve back extension strength, mobility, endurance, and functional disability. Also various exercises, such as lumbar stabilization exercise, motor control exercise, core exercise, lumbar flexion exercise, walking exercise, and bracing exercise, have been proposed to mitigate chronic LBP. Moreover, walking is highly recommended to rehabilitate patients with LBP. It is relatively easy to comply with and is highly cost-effective. It leads to enhanced isometric endurance by increasing muscular endurance and has the potential to eventually prevent LBP (Suh et al, 2019).

The MedX lumbar extension physiotherapy machine (Smart4Health) has been used in both the prevention and treatment of LBP, training the lower back muscles in an isolated manner, leading to an increase of back muscle strength and mass and, in this way, decreasing the effects that LBP will have in the daily life of elderly people (Holmes et al, 1996).

Thus, the suggestion that physical training and low back muscles rehabilitation could contribute positively to the wellness and quality of life of the elderly is within the possibilities to explore in SMART BEAR through locomotor and mechanical metrics.

By designing and creating a synergistic approach targeting LBP and new integrated technological solutions it is possible to contribute for senior citizens compliance to physiotherapy, enhance physiotherapy impact, increase the amount of physical exercise, improve postural stability and, finally, support the wellbeing and quality of life of senior citizens. Moreover, patients suffering from LBP will have a treatment likelihood that will improve their back-muscle strength and mass and decrease their medical care needs.

During the SMART BEAR study, the primary objective with respect to non-specific LBP is to investigate the effectiveness of a multisensory rehabilitation intervention provided via an augmented reality platform combined with exercise versus exercise on the postural control, the physical functioning and the pain intensity in people with non-specific low back pain at baseline assessment and after 6 and 12 months follow up.

Descriptive analysis of SMART BEAR participants with no-specific LBP, will reveal to researchers a more holistic profile of the targeted population, identifying correlations between postural control, functioning, physical activity, events of low back pain and several clinical variables. In order to enlighten the most significant factors among the outcome variables from several prediction models.

Table 2. iKPIs concerning low back pain in the Pilot of Pilots

IKPI ID	DESCRIPTION AND TARGET VALUE	VALIDATION
IKPI 36	Drop-out during MedX physiotherapy training sessions, less than 50% of the total number of participants training in the MedX	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 37	Participants that complied with the physiotherapic MedX sessions will have a decrease in the number of visits to the general practitioner. Participants that completed the MedX training sessions will need 2% less medical appointments (due to LBP) per year	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 38	Improvement of muscle strength and mass (Nm/Kg) for each angle by 15%, when comparing the first MedX session values with the last MedX physiotherapy training session	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 39	Increase of the average number of time spent in good posture by 10% when compared to the number of time spent in bad posture at 6th and 12th month per participant	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 40	Decrease in the VAS* score by 2 points, when comparing the first MedX session VAS value with the last MedX physiotherapy training session VAS value	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 41	Increase in the average daily number of steps	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 42	Increase in the average daily number of steps at 6th and 12th month per participant	Validated in the pilot of the pilots. Reported in D12.5.

IKPI 43	Drop-out during Holobalance training sessions, less than 50% of the total number of participants training in the Holobalance	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 44	Increase in the average daily number of hours of sleep at 6th and 12th month per participant	Validated in the pilot of the pilots. Reported in D12.5.
IKPI 45	Decrease in the number of reported falls at 6th and 12th month when compared to the baseline assessment per participant	Validated in the pilot of the pilots. Reported in D12.5.

*VAS corresponds to the question: "How is your pain level today? Please rate your pain from o to 10 (10 being the worst pain that you ever had and 0 no pain at all)."

3 Technical iKPIs and KPIs

In the following Table 3, the technical achievements obtained are summarised.

Table 3. Objectives and present status

Objective	Status
Continuous participant monitoring, environment sensing, data collection & interoperability	The openHAB (that serves the HomeHab smart-home facility) was finalised and disk images were created that can be replicated. A database service that stores all home sensor values was integrated with the HomeHab. Data aggregation services were created and are accompanied by REST services for integration with the mobile application. The mobile application was extended to a working version. The openHab cloud service was deployed to a private cloud server. The SMART BEAR mobile app collects and tranfers health related data in FHIR format from the connected devices and the provided services to the SMART BEAR platform.
Data analytics, decision making for intelligent and personalised interventions enhancing the well-being and independent living of participants	BDA engine: the following components were assembled to build the BDA engine: Delta Lake, Spark, Trino, Airflow. Also, the workflow for designing, testing, and deploying analytics has been implemented. concerning the design of analytics, some simple outputs, e.g. correlation between parameters, can be obtained, and the work is focusing on more complex ones. General data and statistics are visualised in the dashboard, and also clinical data from the SMART BEAR devices and the synergies with Smart4Health and HoloBalance are shown. Implementation Guide: updated considering also reviewers' recommendations and feedback from Pilot of Pilots (PoP). The information concepts used to facilitate the data analysis phase and the decision-making process have been added. The IG FHIR features resource profiles, extensions, value sets, code systems and examples. Numerically there has been a substantial change. The first version of the IG, dated 09/06/2021, consisted of 51 resource profiles, 2 extensions, 22 sets of values, 8 Code systems, and 60 examples. The third version, released on 03/07/2022, consists of 105 resource profiles, 5 extensions, 37 sets of values, 2 Code systems, and 180 examples. Regarding the non-FHIR repository, the implemented functionality for the generation of the notifications/alerts has been extended to allow the triggering of personalised notifications for the participants and the clinicians for BMI, weight gain and weight loss. Various modifications have been applied to the DSS to allow better performance.
Security, Privacy and Trustworthiness of the platform	Security Component: additional functionality introduced to allow the monitoring of the usage data transmissions (metadata) to the SB@Cloud by those responsible for 1st and 2nd level technical support (role: Help Desk Operator), as well as to fine-tune m2m transmissions triggered by synergies equipment (HoloBalance, Smart4Health). At the same time, the operation (i.e., user management) of the platform were supported. During the PoP the functionality supported by SB@SecurityComponent has been tested in real

Objective	Status
	conditions, during with understanding and fixing bugs resulted in establishing a standardised process of testing, detecting, reporting and repairing technical defects. A complementary (with reference to previous penetration tests) penetration test has been conducted (via a semi-automated method) that again involved the attempted breaching of all SB@Cloud components and associated application systems (e.g., application protocol interfaces (APIs), frontend/backend servers) to uncover vulnerabilities, and to confirm the correction of any weaknesses observed. Minor issues detected have been fixed.
User-centric implementation and integration of the platform	Dashboard: the implementation and mapping of clinical questionnaires was completed. Also the mapping of medical conditions and medications is almost complete in compliance with the HL7-FHIR standard. The following functionalities were added to the dashboard: the patient profiling, which opens automatically the questionnaires and examinations required for specific clinical conditions, and the localization algorithm which selects automatically the language based on the browser settings and supports all the languages of the project. Data aggregation services are created, and their corresponding services are allowing the integration with the mobile application and the data cloud servers. The openHAB cloud services were also integrated with each openHAB device. The integration of the already developed components and the methodology for maintaining the platform were performed and it worked satisfactorily during the PoP. The framework for migrating to Kubernetes is set up as well as the methodology for performing automatic testing.
Demonstrate and validate the use of the SMART BEAR platform at TRL 9 via five large scale pilots, involving 5100 participants across six EU countries	During the third year of the project the Pilot of the Pilots (PoP) continued its activities for the implementation and demonstration of the technical and clinical solutions of SMART BEAR in synergy with Smart4Health and Holobalance projects. Following the pre-recruitment, several dissemination activities were set up among the elderly to assess their willingness to participate. Among the dissemination activities that have contributed more to attract potential participants, we highlight the Madeira Digital Health and Wellbeing actions, in particular the demonstration sessions held in the exhibition stand that attracted 2332 visitors so far. The Madeira Digital Health and Wellbeing stand is also a physical help desk responsible to create a permanent point-of-contact with the PoP participants. During pre-recruitment, the PoP clinical team checked the participants' conditions and clinical criteria for those who were interested in participating and who would benefit from SMART BEAR. By M26, around 100 patients had already been scrutinised. Considering the preconditions of the PoP, which is also to test the synergies with Smart4Health and HOLOBALANCE projects, the pilot needed to set-up the infrastructure to accommodate such projects' solutions. For the participants with balance disorders, a Holobox was installed in Madeira (M25) to perform an augmented virtual reality enhanced rehabilitation

Objective	Status
	training within the synergy with HOLOBALANCE. Additionally, two MedX machines are also available in Madeira for participants to carry out the training on the prevention and treatment of LBP since M21, as part of the synergy with Smart4Health. Recruitment and baseline assessment started on M26 at the premises of Dr. Rui Adriano de Freitas Health Centre and Machico Health Centre, after the local ethics approval and the assessment of the final protocol for SMART BEAR. After the baseline assessment, devices were delivered to the first 12 participants (M27) according to the health conditions and comorbidities assessed by the PoP clinical team, followed by an initial training and personalised explanation of all the operations that can be done with the SB@App and the different devices. The preparation of the devices occurred prior to the handover of the kits (M26), pre-configuring all the devices, installing the necessary apps according to the different combination of comorbidities and pre-pairing the devices. Additionally, on M28 it was performed the home installation of @SB Home Hubs for the first participants with cognitive disorders. From M26 to M36, pre-recruitment, recruitment and assessment were performed within the PoP, in order to increase the number of participants registered in the SMART BEAR platform and its usability. During this time, more than 300 people were approached to participate in SMART BEAR. Dissemination was accomplished through several events (72), presenting SMART BEAR in several elderly institutions and social media, setting up device demonstrations in the Madeira Digital Health stand, showcasing SMART BEAR in day care centres and participating in local TV shows with a large elderly audience. At M36 there are 77 participants registered in the SMART BEAR platform from the 100 patients envisaged in the Italy-Portugal pilot, furthermore, 25 pre-recruited elderly will be recruited in the next few weeks. At the time being, 15 participants already finished the Smart4Health and Holobalance training protocol and continue to actively monitor their health through the SMART BEAR devices. In addition, by M35 the PoP clinical team had already performed the 6th month assessment to **16** participants who already comply with this criterion.
To maximise the project's impact and create a sustainable solution	Submission of D13.2 (D58) "Interim Communication, Dissemination, Impact Creation, Exploitation and Standardisation Report" apart from summarising all relevant activities that have taken place during the best part of the project's third year, it also provides for the basis for a sustainable exploitation schema as all exploitable assets per owner (project beneficiary) have been identified and are currently under monitoring. Project's technical overview along with the current status of the SMART BEAR platform (the version of being used and evaluated during the "Pilot of the Pilots" in Madeira) was also presented at the Forum International de Cybercequritè 2022 in Lille France. SMART BEAR is also contributing to the activities of the following EC LSP WGs:

Objective	Status
	• **WG1 Dissemination**: SMART BEAR was selected as leading partner of the WG1 for 2022, responsible for the coordination of the different tasks to be organized. SMART BEAR partners QUIRON and ANA are co-leading the tasks performed so far: 1. Charing of the monthly meetings; 2. Coordination of the 1st WG1 Newsletter; 3. Organization of the 2nd Open Webinar together with the World Health Organization in relation to Large Scale Pilots, with a focus on assistive technology to support integrated care. • **WG2 Use Cases**: Aim of this cluster is the sharing of experience and the collaboration among LSPs towards the: 1. Creation of a common use case scenario template that could be useful to future researchers in the domain of health, 2. Identification of common areas of scientific interest among LSPs (i.e. multiple LSPs address the same medical conditions , multiple LSPs could make use of the same data management / analysis materials and methods), 3. Design of potential synergies in the context of these common areas and 4. The maximisation of each LSP's impact in the domain of health using the available and not extra resources. • **WG3 KPIs**: SMART BEAR representatives actively participated in the WG3 workshop "Evaluation and exploitation in large scale pilots: lessons learned by Platform Uptake and Open DEI'' held online on the 19th of May 2022. Beyond being a good occasion to catch up and exchange the progress made in evaluation and selection of KPIs among all the Cluster's members, this workshop laid out the future steps for the business impact assessment of the OPEN DEI projects' ecosystem, developed within the frame of OPEN DEI 's Task Force 4 activities. Namely, a Business Impact Assessment Survey was developed and shared with the LSP's. SMART BEAR contributed with its feedback from the exploitation manager during July'22. The results of the survey will be analysed within OPEN DEI activities to map the positioning of ecosystem projects in terms of business impact and provide insights about potential pathways for a better exploitation of the results. The results of the survey and further analysis will be communicated through a webinar by OPEN DEI in September-October 2022. • **WG4 Architecture**: SMART BEAR representatives were deeply involved in the health systems architecture design experience exchange. The collaboration was related to the organisation of the data flows in order to support reliable and secure medical data gathering, storage and analysis. As a result, the common ground was found to make the architecture more universal as a whole and its components reusable in particular. Additionally, the collective approach to description and comparison of SMART BEAR similar systems was developed. The active participation of the SMART BEAR team in the LSP resulted in the publication of the article "Reference Architectures, Platforms, and Pilots for European Smart and Healthy Living Analysis and Comparison". • **WG5 GDPR:** SMART BEAR partners have participated in the exchange of knowledge on consent with other member states that participate in projects similar to SMART BEAR, in order to find out which is the best protocol to follow to obtain the consent. Some of the achievements of

Objective	Status
	the Health Cluster have been: i) OPEN DEI Healthcare cluster, participated by Smart Bear representatives, has obtained a CEN Workshop Agreement (CWA). The CWA will be the basis to define a guideline for obtaining the most suitable consent for the use of digital health innovations throughout European Health Data Space.https://www.cencenelec.eu/media/CEN-CENELEC/News/Workshops/2022/2022-06-13%20-%20Digital%20Health/draftprojectplan.pdf; ii) EU DATA STRATEGY Public consultation OPEN DEI - EHTEL response. SMART BEAR's technical dimension and the platform (the version being used and evaluated during the PoP) were presented at the Forum International de Cybercequritè 2022 in Lille France (July 2022). The already defined architectural components and interfaces for the m2m transmission of health data (FHIR compatible) with Smart4Health and HoloBalance solutions have been fully implemented and tested within the PoP in Madeira. Additionally, synergies with other projects are channeled through SmartBear's leadership of EC LSP WG1 on Dissemination hence allowing for reaching out to a broader audience of both academic / scientific nature, as well as communities of citizens in need of healthcare and assistance.

3.1 Technical Milestones

The following Table 4 lists the ten milestones (MS10 milestone was added in AMD-857172-15) of the Smart Bear project, along with their description and actual status.

Table 4. Milestones and present status

Milestone number	WP number	Description	Status
MS1	2	Final architecture of the SMART BEAR platform	**Achieved on month 10**, with the release of D2.2
MS2	3	Initial version of the SMART BEAR @ Home enabling components	**Achieved on month 17**, with the release of D3.1.
MS10	3, 4, 5, 6	Pilot of the Pilots prototype delivery and demonstration	**Achieved on month 21**, with the release of D3.2, D4.2, D5.2 and D6.7.
MS3	3, 4, 5, 6	Initial version of tools, models and integrated platform	**Achieved on month 30**, with the release of D6.1.
MS4	6, 7, 8, 9, 10, 11	Initial prototype demonstration and pilot validation	Ongoing, with D6.2 and D8.5 already submitted. It will be achieved with the release of D7.1, D8.1, D9.1, D10.1 and D11.1.
MS5	12	1st Cross-Pilot Evaluation	Ongoing, will be achieved with the release of D12.3.

MS6	6, 7, 8, 9, 10, 11	Second prototype demonstration and pilot validation	Ongoing, will be achieved with the release of D6.3, D6.4, D7.2, D8.2, D9.2, D10.2 and D11.2.
MS8	12	2nd Cross-Pilot Evaluation	Ongoing, will be achieved with the release of D12.4.
MS7	6, 7, 8, 9, 10, 11	Final prototype delivery and demonstration	Ongoing, will be achieved with the release of D6.5, D6.6, D7.3, D8.3, D9.3, D10.3 and D11.3.
MS9	12	Cross-Pilot validation & Final Evaluation	Ongoing, will be achieved with the release of D12.5.

3.2 Technical KPIs: Progress

The following Table 5 lists the KPIs associated with the technical work and achievements.

Table 5. Summary of KPIs and current status

KPI	Tech WP	Description	Results
KPI-1	WP3	The monitoring solution will comprise at least 5 heterogeneous devices (smart home devices/sen- sors/hubs, wearable devices, smartphones, and medical devices like hearing aids, blood glucose meters, as described in Table 2, to effectively demonstrate its general applicability.	• Platform FHIR ready, Transmition ready, • MVP (supporting the PoP): integration of 6 devices and 2 sensors • List of devices/sensors has been finalised • HAs integration is **still in progress** • Results to be validated in pilots
KPI-2		The monitoring solution will increase accuracy of state identification at least 25% with respect to each of the individual devices when that device is used in isolation.	
KPI-3		The monitoring solution will help keep drop-out rates due to various reasons (interoperability, usability, data security/privacy issues, user acceptance) low (<5%).	
KPI-4	WP3-4	Deliver basic analytics to infer high level parameters regarding the (1) functional, (2) physiological, and (3) behavioural state of each participant, as well as their (4) safety, (5) risk, (6) social interactions, (7) cognitive, (8) sensory, and (9) balance profiles.	• BDA integration completed • Several descriptive analytics and 1 ML (blood preasure) have been integrated, based on PHR, available device usage data, SH4 and HB data
KPI-5		The basic analytics will improve the comprehension of individual health	

		status and condition by 20% compared to using raw sensor data.	
KPI-6	WP4	Deliver advanced data analytics and learning capabilities for large scale analysis of the comprehensive datasets collected from all pilots and participants, which are capable of dealing with the volume and velocity of data collected by the various devices used for monitoring participants.	• PoP: Dashboard supports end-user management, patients management, Analytics, Interventions • Descriptive and ML analytics considering PHR, questionaries answers, usage data from the first set of devices/sensors, SH4, HB raw/aggr data • Analytics results to trigger technical (non-medical) messages/alerts (automatically), Smart hints to CCMs
KPI-7		The advanced data analytics and learning capabilities will produce evidence enhancing existing health delivery services for the targeted 5 health conditions and their related comorbidities, by reducing their overall delivery cost by at least 20%.	
KPI-8		The advanced data analytics and learning capabilities will provide evidence useful to at least 4 types of stakeholders (HCP, businesses, national governments, professional societies).	
KPI-9	WP6	The decision support capabilities will cover at least 5 medical areas: (i) hearing loss, (ii) cardiovascular diseases, (iii) cognitive impairments, (iv) mental health issues, and (v) balance disorders.	• PoP: 77 participants, (common with S4H - HB: 15). • Completed: M2m digestion of Smart4Health and HoloBalance data structures (for common study participants)
KPI-10		The decision support capabilities will reduce unnecessary interventions by at least 20% in hearing loss and balance disorders, and 10% in CVDs, cognitive impairments, and mental health issues.	
KPI-11		The decision support capabilities will speed-up the delivery of appropriate medical interventions by healthcare providers by at least 20% of the current time needed.	
KPI-12	WP3	Deliver applications for participant screening and monitoring, and delivery of the *SMART BEAR* platform interventions.	• Profiling is supported • Rule-based personalised interventions (in D3.1) • Mobile app functionality enhancements (testing phase): questionnaires, Medication app,

			OTAGO app, Diet App, Serious games, multilingualism
KPI-13		Acceptance of the application interventions by the users, reporting a score higher than 80 in Software Usability Scale (SUS) and higher than 80% in Technology Assessment Methodology (TAM).	
KPI-14	WP5	The privacy-preserving and secure by design data handling capabilities delivered will ensure the key properties of: (i) anonymity, (ii) confidentiality, (iii) privacy, and (iv) integrity of data.	• PoP: final version of the SC functionality SC been tested, supporting RBAC, token-based access, pseudonymisation, personal data IDs and devices management • DPA has been signed (ICCS, host org)
KPI-15	WP5	The privacy-preserving and secure by design data handling capabilities delivered will ensure the afore-mentioned key properties for data: (i) in transit, (ii) at-rest, and (iii) in processing.	• 4 Security assessments conducted (via semiautomated fashion): potential vulnerabilities were identified and resolved • S/P platform integration **still in progress**
KPI-16		The privacy-preserving and secure by design data handling capabilities delivered will have demonstrable security monitoring policy compliance and auditability for all properties contained in respective security policies.	
KPI-17		The privacy-preserving and secure by design data handling capabilities delivered will be compliant with the encryption guidelines of NIST.	
KPI-18		The security framework will respect all ISO/IEC 27000 family of standards and also be able to operate in different security domains for various identifications, by providing application-specific certifications and license requirements.	
KPI-19	WP6	The integrated smart adaptive living solution will interoperate with existing open-source smart platforms, specifically with FIWARE.	• PoP: MVP version of the integrated platform currently been tested (all components of Production instance) • CI/CD, Ticketing: feedback is gathered, testing procedures in place
KPI-20	WP3	The integrated smart adaptive living solution will integrate sensing and actuation mechanisms such as these described in Table 2.	• PoP: MVP (supporting the PoP): integration of 6 devices and 2 sensors

			• List of devices/sensors has been finalised • HAs integration **still in progress** • Results to be validated in pilots
KPI-21	WP6	The integrated smart adaptive living solution will provide interfaces to the platform for at least 4 types of people: (i) elderly users, (ii) their caretakers, (iii) health-care professionals, and (iv) health-care managers & business/regional decision makers.	• First version of supplementary material for installation, deployment and usage of the technical solution
KPI-22	WP2	At least 5 individuals from each pilot site will participate during the design of the user-centric platform (either in focus groups, individual interviews, etc.).	• Platform currently been used within the context of PoP • Usability evaluation results to be presented in D12.4, D12.5
KPI-23		At least 90% of the participants will evaluate the final SMART BEAR platform usability as greater than or equal to 8 in a range of 1-10 in at least 80% of the questions in the evaluation questionnaires.	
KPI-24		At least 95% of the participants will be using the SMART BEAR platform for the whole duration of the pilot, indicating a drop-out rate of less than 5%.	
KPI-25		At least 90% of the participants will be willing to recommend the SMART BEAR platform to a friend by the end of the project	
KPI-26		The integrated smart adaptive living solution will improve the effectiveness of personalized care by at least 30% compared to what is currently achievable.	
KPI-27		The integrated smart adaptive living solution will reduce visits to healthcare providers by at least 40%.	
KPI-28		Delivery and demonstration of the TRL-9 integrated SMART BEAR platform to each of the partners supporting the pilots.	
KPI-29		At the end of the first three months of a pilot, metrics will be collected for each individual, according to their health conditions and their related comorbidities. These metrics should be improved by at least 10% by the end of the pilot.	

KPI-30		At the end of the first three months of a pilot, metrics will be collected for all individuals in the pilot site, according to their health conditions and their related comorbidities. The metrics should be improved by at least 10% by the end of the pilot.	
KPI-31		At the end of the first three months of a pilot, metrics will be collected across all pilot sites. These metrics should be improved by at least 10% by the end of the pilot.	
KPI-32		Achieve the project's dissemination targets as defined in Table 11 of Sect 2.2.6.	
KPI-33		Contribute to standards, such as CEN ISO 13606 EHR-communication standard, and to the series of standards known as Concurrent-Use-Standards (CEN: EHR-communication, Health Information Service Architecture and System of Concepts for Continuity of Care) - see Sect. 2.2.3.	
KPI-34		At least 2 clustering/concentration activities with the CSA supporting the activities defined under "DT- ICT-13-2019: Digital Platforms/Pilots Horizontal Activities". The activities will take place at M03 before the design of the solution and then at M38 in order to report on the initial results of the pilots.	

4 MAFEIP Evaluation framework

MAFEIP[1] (Monitoring and Assessment Framework for the European Innovation Partnership on Active and Healthy Ageing) is a monitoring framework, developed by the European Community to specifically support the continuous evaluation of ICT-enabled social and health innovations, such as new care pathways, devices, surgical techniques, organisational models, etc. The MAFEIP tool performs a simulation that measures the likelihood that the assessed interventions will achieve their expected impacts in terms of both increased efficiency and improved health and quality of life of the beneficiaries. In this way, it is possible to estimate the effectiveness in healthcare resource use, societal resource use and health-related quality of life that result from using the considered innovation, instead of current care practises.

The MAFEIP simulation is based on the principles of Decision Analytic Modelling (DAM) through Markov models, which simulate the transition of the target population between its modelled states. The tool includes 3 different models, which differ in the number of their states (3, 4 or 5, as shown in the next Figure 4-1), allowing to select the model that better adapts to the intervention to simulate. The states of the model represent the Baseline state, Disease/impairment or Progressive disease stages of the condition that the intervention aims to prevent, relieve or cure, and Dead stage. The Markov model is completely defined by *the initial distribution among the states*, *the transition probabilities* and the *parameters associated with each state*. Moreover, each state of the model is also defined by an amount of resource use (costs) and quality of life (utility or health outcomes), which respectively represent the average resource use and quality of life of a patient in that health state. Thus, an implementation of the MAFEIP model requires the input data that allows to define these parameters for each state of the model, as well as to simulate the transitions of the population among the various stages.

The simulation considers a cohort of patients (which in our case are the patients enrolled into the various pilot studies) defining an initial distribution among the starting states. Then, new data is acquired during a cycle, corresponding to the time frames where the patient have been observed and cured with the proposed innovation. In this way, the appropriate transition probabilities are applied and the distribution of patients in each state is adjusted following the real distribution of the cohort in the pilot studies. As the model runs over a specified time horizon (i.e., several cycles), it is possible to estimate the incremental costs and outcomes associated with a particular intervention, based on the data introduced into the tool during specific time frames, by analysing the MAFEIP simulation's results.

[1] https://www.mafeip.eu/

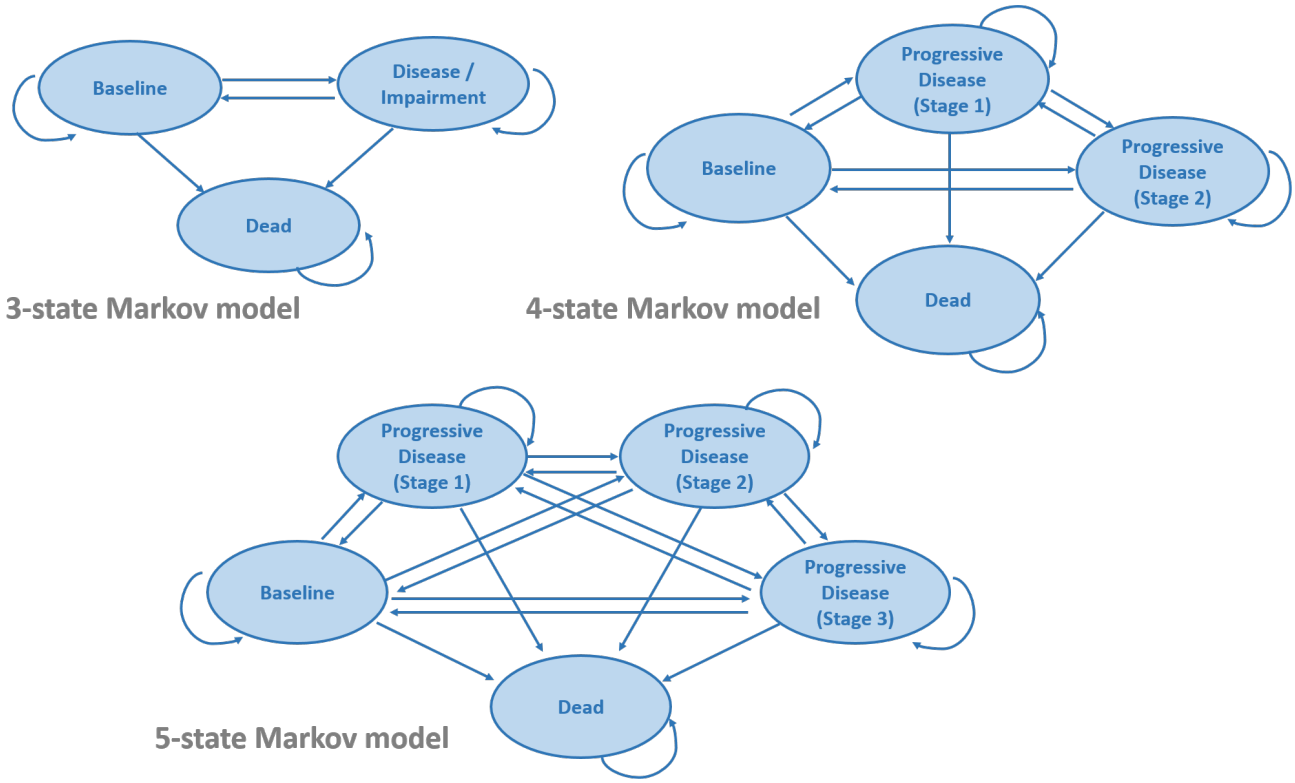

Figure 1. MAFEIP Markov Models

As described above, the implementation of a MAFEIP simulation requires defining the model and its states, providing the tool with all the needed data, following a set of predefined forms. Therefore, during the initial phases of the WP12, several meetings and discussions have been held between the medical and technical partners of the project, with the purpose of developing an implementation plan of the MAFEIP framework for SMART BEAR. The main purposes of the implementation plan are to model the stages of the diseases and identify the inputs required by the MAFEIP framework to perform its simulations. In detail, the implementation plan identified the models and the respective numbers of the states (in some cases, their specific meaning, adapting it to the considered disease/condition) and the input data, as well as any other additional required data, the most effective ways to acquire the inputs, the time frames of the simulations and others. In this way, it will be possible to implement a MAFEIP simulation for each disorder monitored through the SMART BEAR platform in the various pilot studies, including the Pilot of Pilots case (introduced during the development of the project). The result of these simulations will provide a further assessment of the obtained results and, in general, of the proposed platform. In the following of this Section, all the details of the implementation plan will be described.

4.1 Models and monitored parameters

The MAFEIP tool provides a web-based interface for the definition of the model and to input the required data, supporting in this way the implementation of the model by defining and requesting the data in different steps. In the preliminary step of the process, the number of states of the Markov model must be selected. As shown in the previous Figure 4-1, the number of these states depends on the progressive stages of the considered disease/impairment. Thus, the medical teams involved in the project had to identify for each disease tested in the pilots their corresponding possible

progressive stages, allowing in this way to set the number of states of the Markov model. Moreover, as detailed in the following of this Section, they also had to identify the ranges of values of the parameters monitored through SMART BEAR corresponding of each state of the model, in order to obtain the data required by the tool to calculate the transition probabilities. The following Tables show the various stages identified and the corresponding ranges of the monitored parameters. In the following tables, it is possible to note the monitored parameter(s), the corresponding ranges that are mapped into a state of the MAFEIP model and, if necessary, the meaning and/or the description of the stages.

Table 6. Mild Cognitive Disorders parameters for MAFEIP models

MILD COGNITIVE DISORDERS			
Parameter: MOCA Score			
Parameter Ranges	***Description***	***Notes***	**MAFEIP Stages**
> 26/30*	**Cognitively normal**	Healthy	**Baseline**
18-26/30*	**MCI**	Pre-dementia	**Stage 1**
11-17/30*	**Mild Dementia**	Dementia stages	**Stage 2**
6-10/30*	**Moderate Dementia**		
<6/30* or not testable	**Severe Dementia**		**Stage 3**
			Death

* These scores are corrected by age, education and cultural background as recommended by DSM-V and ICD-12.

Table 7. First group of CVD parameters for MAFEIP models (Hypertension, Coronary Artery Disease, Arrhythmias)

CVD - Hypertension, Coronary Artery Disease, Arrhytmias											
Parameter: visits to emergency department due to hypertension urgency/emergency per year				Parameter: time in therapeutic range (BP values)				Parameter: Heart Score			
Parameter Ranges	*Description*	*Notes*	*MAFEIP Stages*	*Parameter Ranges*	*Description*	*Notes*	*MAFEIP Stages*	*Parameter Ranges*	*Description*	*Notes*	*MAFEIP Stages*
N0 (12 months prior to recruitment)			**Baseline**	**>70%**	**percentage of BP values within the personalised levels**		**Baseline**	**HeartScore0 (HS0)**	**Heart score change in 6, 12 mths**		**Baseline**
N = N0 + 2			**Stage 1**	**50-70%**			**Stage 1**	**HS > HS0**			**Stage 1**
N= N0 + 4			**Stage 2**	**30-50%**			**Stage 2**				**Stage 2**
N > N0 + 4			**Stage 3**	**<30%**			**Stage 3**				**Stage 3**
N/a	n/a	n/a	**Death**	n/a	n/a		**Death**	n/a	n/a		**Death**

Table 8. Second group of CVD parameters for MAFEIP models (Hypertension, Coronary Artery Disease, Arrhythmias)

CVD - Hypertension, Coronary Artery Disease, Arrhythmias											
Parameter: hypertension mediated organ damage: hypertension related retinal disease, chronic kidney disease, coronary artery disease, left ventricular hypertrophy, peripheral arterial disease, aortic disease				**Parameter: physical capacity**				**Parameter: BMI**			
Parameter Ranges	***Description***	***Notes***	***MAFEIP Stages***	***Parameter Ranges***	***Description***	***Notes***	***MAFEIP Stages***	***Parameter Ranges***	***Description***	***Notes***	**MAFEIP Stages**
N0	all above conditions will be checked if present on recruitment and rechecked in 6, 12 mths		**Baseline**	minutes in physical activity, > 150 per week			**Baseline**	18.5-24.9			**Baseline**
N > N0			**Stage 1**	100-150			**Stage 1**	25-29.9			**Stage 1**
N>N0			**Stage 2**	50-100			**Stage 2**	30-34.9			**Stage 2**
myocardial infarction, Stroke			**Stage 3**	<50			**Stage 3**	>35			**Stage 3**
hypertensive emergency leading to death			**Death**	none			**Death**	n/a			**Death**

Table 9. Second group of CVD parameters for MAFEIP models (Heart Failure)

CVD Heart Failure							
Parameter: physical capacity				**Parameter: number of acute HF decompensations**			
Parameter Ranges	***Description***	***Notes***	***MAFEIP Stages***	***Parameter Ranges***	***Description***	***Notes***	**MAFEIP Stages**
NYHA class I	**NYHA check during recruitment, 6mths, 12 moths- NYHA IV are excluded on recruitment based on exclusion criteria**		**Baseline**	N0	acute HF decomp 12 months prior to recruitment		**Baseline**
NYHA class II			**Stage 1**	N =/> N0+2			**Stage 1**
NYHA class III			**Stage 2**	N =/> N0+3			**Stage 2**
NYHA class IV			**Stage 3**	N =/> N0 +4			**Stage 3**
death			**Death**	death			**Death**
			Death				

Table 10. Frailty parameters for MAFEIP model

Frailty			
Parameter: EFS Score			
Parameter Ranges	*Description*	*Notes*	MAFEIP Stages
6-11	Inclusion scores	Only this range of scores will be included	Baseline
≤ 5	Non-frail	Healthy	Stage 1
6-11	Mild Frailty	Apparent vulnerability	Stage 2
12-17	Severe Frailty	Severely vulnerable	Stage 3
			Death

Table 11. Mild Depression parameters for MAFEIP model

MILD Depression							
Parameter: PHQ-9				**Parameter: SQS**			
Parameter Ranges	*Description*	*Notes*	*MAFEIP Stages*	*Parameter Ranges*	*Description*	*Notes*	MAFEIP Stages
0-4	None or minimal symptoms of depression		Baseline	10	Excellent quality of sleep		Baseline
5-14	Mild to Moderate Depression		Stage 1	9-4	Good to Fair quality of sleep		Stage 1
15-19	Moderately severe Depresion		Stage 2	3-1	Poor quality of sleep		Stage 2
20-27 or question 9>0	Severe Depression		Stage 3	0	Terrible quality of sleep		Stage 3
			Death				Death

Table 12. Low back pain parameters for MAFEIP model

Low back pain			
Parameter: Duration of symptoms			
Parameter Ranges	*Description*	*Notes*	*MAFEIP Stages*
0 - 6 weeks	acute	non eligible for the study	Baseline
6 -12 weeks	sub-acute	non eligible for the study	Stage 1
more than 12 weeks	chronic	eligible for the study	Stage 2
			Stage 3
			Death

Table 13. Balance Disorders parameters for MAFEIP model

Balance Disorders (concern of participants regarding falling)							
Parameter: Falls Efficacy Scale – International (FES-I)				**Parameter: Functional Gait Assessment**			
Parameter Ranges	***Description***	***Notes***	***MAFEIP Stages***	***Parameter Ranges***	***Description***	***Notes***	***MAFEIP Stages***
0 -16	**no concern about falls**	not included in the study	**Baseline**	**30 - 22**	**values opposite to risk of falling**	recruitment in balance rehabilitation group (Holobalance or OTAGO)	**Baseline**
16 -19	**low concern about falls**	recruitment in balance rehabilitation group (Holobalance or OTAGO)	**Stage 1**	**below or equal to 22/30**	**risk of fall**	recruitment in balance rehabilitation group (Holobalance or OTAGO) with respect to FES-I values	**Stage 1**
20 - 27	**moderate concern about falls**	recruitment in balance rehabilitation group (Holobalance or OTAGO)	**Stage 2**	**below or equal to 20/30**	**optimal to predict an unintentional fall in the next 6 months**	recruitment in balance rehabilitation group (Holobalance or OTAGO)	**Stage 2**
28-64	**high concern about falls**	reqruitment in balance rehabilitaion group (Holobalance or OTAGO)	**Stage 3**				**Stage 3**
			Death *				**Death ***

Table 14. Hearing Loss parameters for MAFEIP model

Hearing Loss			
Parameter: HA benefit according to GHABP (HA benefit related questions and score)			
Parameter Ranges	***Description***	***Notes***	***MAFEIP Stages***
4-5	**Perfect**	Optimal usage and benefit	**Perfect**
3-4	**Great**	Acceptable benefit	**Great**
2-3	**Moderate**	Acceptable benefit	**Moderate**
1-2	**Slight**	Little benefit of HA usage, to improve	**Slight**
0-1	None	No satisfaction or drop out	**No benefit of hearing aid usage OR no usage at all (DROP OUT)**

The parameters and the identified ranges shown in Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, Table 9, Table 10 and Table 11 will allow to define the number of the states of each model of MAFEIP. It is worth note that the MAFEIP tool can produce a single simulation for each disease, based on the monitored parameters, which must be used as input to automatically calculate the transition probabilities of the stages of the models. The acquisition of these values will be fully automatic, leveraging the BDA engine of the SMART BEAR platform. In this case, no additional effort by the physicians and medical personnel involved in the pilot study will be required.

4.2 General data

The first step of the MAFEIP model design (after the preliminary step described above) foresees the input of the general data related to the intervention. In this step, the user must specify the

characteristics and the aims of the intervention that has to be evaluated. This information is divided in the following four subsections, which are:

- *General information;*
- *Setting and target population;*
- *About your intervention;*
- *Evidence*.

In the *General information* subsection, only the name of the intervention and its action groups, as defined in the Strategic Implementation Plan of the EIP (European Innovation Partnership) on AHA (Active and Healthy Aging)[2] must be provided. The *Setting and target population* includes the demographic and geographic characteristics of the target group, the condition that the intervention aims to prevent or cure and the typical characteristics of the target group (e.g., the comorbidities). An example of such kind of data is shown in the next Figure 4-2. The *About your intervention* subsection requires to select if the intervention relates to a device, a protocol/treatment method, or other, as well as a brief description of the clinical implementation of the intervention and the current care situation without the intervention; finally, it requires to select the stage of the development of the intervention among the pre-trial, clinical trials ongoing, pilot implementation, routine use and other. An example of this subsection is depicted in the next figure. The last subsection of the first step, *Evidence*, collects some information related to how evidences of the study are collected, such es, the type of study, the presence of a control group and the adoption of evident on the impact on Health Related Quality of Life (HRQoL), as shown in the subsequent figure. It is important to underline that all the aforementioned data required in the subsections of the first step of the MAFEIP model design can be easily obtained with the support of the partners involved in the pilots and must be provided only once, at the moment of the definition of the models.

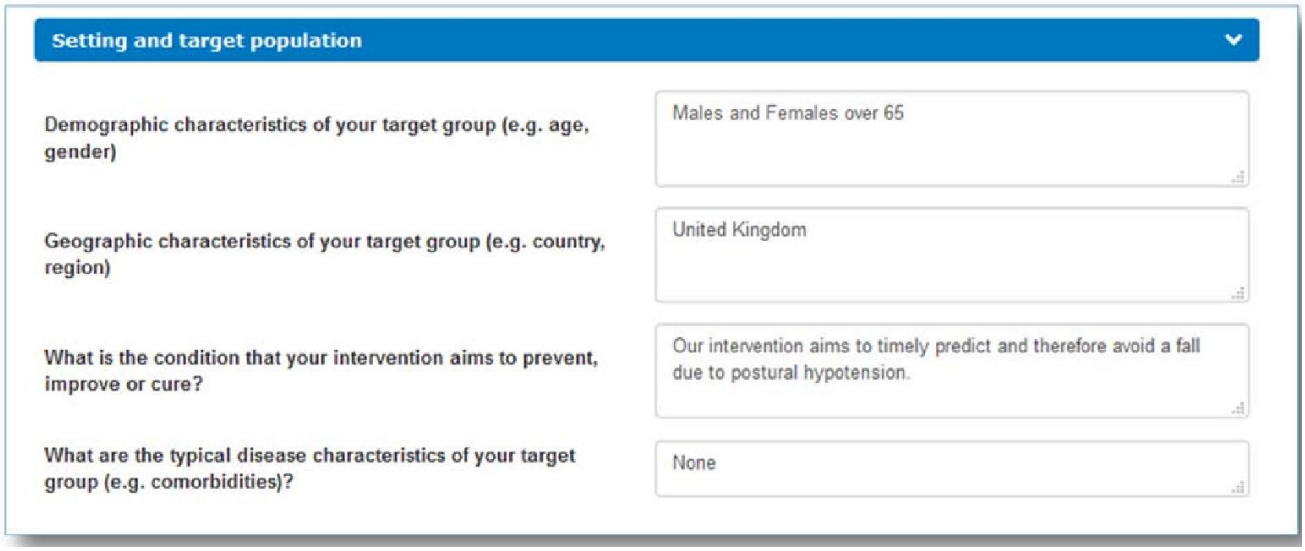

Figure 2. Example of "Setting and target population" data

[2] The plan and the action groups are available at https://futurium.ec.europa.eu/en/active-and-healthy-living-digital-world

About your intervention

Your intervention relates to a:	(●) Device ( ) Protocol/treatment method ( ) Other, please specify
Brief description of the clinical implementation of your intervention (e.g. who uses it, when, in which setting, how often etc.)	Elderly people over 65 years old who are living at home, in nursing homes, or at the hospital will be equipped with a device that would measure Heart Rate Variability (used to predict possible postural hypotension and subsequent falls).
Brief description of the current care situation without the intervention (e.g. current treatments provided, whether the intervention replaces or complements the current therapy, etc.)	There is currently no other treatment or such device.
Does your intervention have an impact on health and resource use compared to current care? (please describe briefly).	1/3 of people over 65 and 1/2 of people over 80 fall at least once a year. 30% of the falls are due to postural hypotension. Falls cost the NHS more than 2 billion GBP a year and with the number of people over 65 increasing, these costs may also rise by 2020. Our intervention would like to decrease these numbers.
Stage of development of your intervention	(●) Pre-trial (e.g. proof of concept) ( ) Clinical trials ongoing ( ) Pilot implementation ( ) Routine use ( ) Other, please specify

Figure 3. Example of "About your intervention" subsection

Do you collect empirical evidence on the *effectiveness* of your intervention?	(●) Yes ( ) No
What is the study design with which evidence on effectiveness is gathered?	( ) Randomized Controlled Trial ( ) Longtitudinal observational ( ) Cross-sectional observational ( ) Case-control ( ) Other, please specify
Was there a control group against which the effectiveness was assessed?	(●) Yes ( ) No
Number of patients in intervention arm of study	
Number of patients in control arm of study	
Do you collect empirical evidence on the impact of your intervention on *resource use*?	(●) Yes ( ) No
What is the study design with which evidence on resource use is gathered?	( ) Piggyback study with clinical trial ( ) Micro-costing study ( ) Patient records based study ( ) Other, please specify
Was there a control group against which the effectiveness was assessed?	(●) Yes ( ) No
Number of patients in intervention arm of study	
Number of patients in control arm of study	
Do you collect empirical evidence on the impact of your intervention on *Health Related Quality of Life (HRQoL)*?	(●) Yes ( ) No
What is the study design with which evidence on HRQoL is gathered?	( ) Randomized Controlled Trial ( ) Longtitudinal observational ( ) Cross-sectional observational ( ) Case-control ( ) Other, please specify
Was there a conrol group against which the effectiveness was assessed?	(●) Yes ( ) No
Number of patients in intervention arm of study	
Number of patients in control arm of study	

Figure 4. Example of "Evidence" subsection

4.3 Model Input data

The second step of the MAFEIP model design is devoted to the model input data, and it has four different subsections:

- *Set-up;*
- *Probabilities;*
- *Costs;*
- *Utilities* (also called *quality of life weights*).

The required data mainly comes from the pilot studies, as it will be shown more in detail in the next subsections. When required, the SMART BEAR Big Data engine will implement specific analytics, able to automatically provide the MAFEIP model with the data. The only drawback, due to the closed software implemented in the MAFEIP framework, is that it is not possible to directly connect it with the BDA engine of our platform, allowing to automatically feed the various forms in the MAFEIP web application for the definition of the models: the obtained data must be manually inserted, although this work has to be done only one or two times per year, for each implemented MAFEIP model.

4.3.1 Set-up

The *Set-up* subsection requires some basic info related to the main parameters of the simulation that will be performed. It first asks the discount factor for costs and utilities (a standard 3,5% discount rate is set by default and can be adjusted if required – in the MAFEIP SMART BEAR implementation the standard default value will be used), which will be used by the tool to evaluate the future costs at today's equivalent. It also needs the time horizon for the analysis in year, in order to run a number of cycles of the models, equal to the years of the time horizon. In the project's implementation, more simulations with different time horizons will be performed, allowing to evaluate both short-term and long-term effects of the proposed interventions. The data related to target population must be also specified, providing the minimum and the maximum age of the patients, their gender (male, female or both), the county and the currency. The simulated model can run the simulation only for one specific age-gender combination and so also the gender and the age to simulate must be finally indicated. The SMART BEAR models will run the model for both genders and for several average ages of the patients (depending on the average ages of the patients enrolled in the pilots). This information will be obtained (in an anonymised way) from the BDA engine of the platform.

4.3.2 Probabilities

The *Probabilities* subsection of the model allows to input data on the probability of an individual to move from each state defined in the model to the others. The data must relate to a Control Group (patients being treated with the standard care – if a control group in included in the pilot to simulate) and to the Intervention Group (patients receiving the proposed intervention). In detail, the model implementation requires the initial distribution among the states of the model, namely the proportion (in percentage) of the patients that respectively are in the baseline state and in the disease or progressive disease stages (following the model stages previously defined in the preliminary step and, in the SMART BEAR case, previously shown in Table 3 to Table 11). The transition probabilities must be provided in both incidence and recovery cases, indicating the percentage of patients that respectively move from a better stage to a worst one of their condition (baseline to progressive disease stage 1, disease stage 1 to disease stage 2, etc.) and move to an improved stage of the condition (disease stage 2 to disease stage 1, disease stage 1 to baseline, etc.). These percentages will

be obtained from the real data observed during the pilot study, in a six months horizon, extracting them through specifically developed analytics of the BDA engine of the platform.

The transitions to the dead stages will be automatically extracted from the Human Mortality Database by MAFEIP, associating the data related with the cohort under test. This rate, if required, can also be adjusted by specifying a Relative Risk (RR), which estimates the mortality in a certain population (e.g., people who participated in the study) divided by the mortality in a reference population or condition (in this case, from the Human Mortality Database). The tool also allows to directly input the mortality rates associated with the cohort under assessment, if data on mortality rates associated to the relevant study cohort are available.

4.3.3 Cost inputs

The cost inputs required to simulate in MAFEIP framework the cost-effectiveness of the intervention includes both the intervention and the standard care scenario costs. The costs are divided into two main classes: *One off and annual recurrent costs* and *Health state costs*. The *One off costs* are the cost per patient included at the implementation point (and occurred only once). In this case, the SMART BEAR MAFEIP implementation will include the costs of the devices used to monitor the diseases, obtained from the respective procurements, as well additional one-off costs, required to implement and set-up the pilots and enrol the patients, which will be estimated in average by each pilot site, depending these costs on the country and the considered disease/pilot study.

The *recurrent cost per year* represents the annual recurrent cost incurred for delivering the intervention and consists of two components: the variable costs of the intervention that are incurred for each year (e.g., annual software licence, maintenance, etc.) and the share of annual fixed costs per patients currently treated by the intervention (e.g., the annual cost of the infrastructure used for all patients, divided by the number of patients treated by the intervention). In our implementation plan, the first part of these costs will be evaluated taking into account the annual costs for each intervention in each disease/pilot (e.g., email services subscriptions, maintenance services, etc.), while the second part will be evaluated by each pilot site.

The *Health state costs* includes the *healthcare costs* and the *societal costs*. The first one represents the costs for the intervention in each state of the disease and are defined per person. In detail, these costs refer to the resource use within the healthcare system, excluding the cost for the intervention itself. Examples are hospital inpatient episodes, primary care visits, and stays in nursing homes, etc. The latter one includes the related costs outside the healthcare sector, such as out-of-pocket payments, travel costs for patient, productivity losses, costs of relatives acting as caretakers at home and others. In these cases, each pilot sites will provide an average estimation of both costs, following the sample calculations for Baseline and Progressive Disease Healthcare and Societal Costs tables provided by the MAFEIP guidelines.

4.4 Simulation

After the definition of the models for each disease and/or pilot of the SMART BEAR project, following the approaches previously described, it will be possible to run the simulations. The model output section of the MAFEIP tool allows to show the results of healthcare or societal cost perspective: both analyses will be performed and shown in the final report of the MAFEIP-based assessment of the SMART BEAR project. Moreover, as already mentioned, several simulations will be performed, considering the various diseases, genders, average ages of the patients, as well as different time horizons.

5 Usability Evaluation Framework

The Usability Evaluation Framework describes the metrics and procedures for evaluating both the SB app and dashboard. The aim is to analyze usage, acceptance and usability by end-user and to provide insights for the in-progress process of redesign and optimization. The plan was defined by considering:

- recommendations and comments from the General Project Review Consolidated Report (Ref. Ares(2022)3019849 - 13/04/2022), in the specific the need to assess SB technology by using standardized tools and relying on both qualitative feedbacks and quantitative data;
- KPIs related to Usability, Acceptance, and Usage of the SB technology (KPI-13, KPI-23, KPI-24, KPI-25, iKPI 21);
- the specificity of the application domain and the typology of the user (Participant, Clinician).

To find the most suitable standardized tools, a literature review was conducted on PubMed, Scopus, IEEE Xplore, and Google Scholar by using the query:

(evaluation OR questionnaire OR evaluation tools OR quantitative evaluation) AND (platform OR dashboard OR mobile application OR app) AND (mHealth OR m-health OR eHealth OR e-health) AND (usability OR acceptance OR satisfaction OR usage).

Review papers were included in the research for their necessary contribution to get a comprehensive insight into the state of the art) Hajesmaeel-Gohari et al., 2022). A number of 57 articles were retrieved, of which 5 were systematic reviews (4 domain-specific and 1 generic). The most exploited tools and measures are summarized in Section 7.1, Table 1. The most frequently applied definition of Usability is the (ISO 9241-11).

Among the identified tools, a selection was made based on tools: 1) containing a small number of items (to avoid an overloading of the evaluators with too many questions considering preliminary feedback from PoP), 2) including meaningful categories 3) most widely adopted and standardized.

The analysis was carried out with particular attention to the specificity of the tool for the health context, since, as discussed in the literature [2], general instruments are not sufficient and it is important to consider also health-specific attributes of Usability. As an example, from the point of view of a patient it is important to assess if health information and progress are well presented and understood, while from the point of view of the clinician it is important to assess if it is simple to monitor health conditions of the patients and to fill-in/update information.

5.1 Methodology

The adopted approach is a mixed quantitative-qualitative evaluation methodology, based on the administration of standard tools, open-ended questions, and the analysis of data extracted from log files.

Evaluation subjects are IT experts and Clinicians from the SB consortium, and Participants recruited in the project. The IT experts will be involved in the evaluation of both the SB app and the SB dashboard, the Clinicians will evaluate the usability of the SB dashboard, and the Participants will be asked to evaluate the SB app (see Figure 7).

The evaluation goes through two stages that will be carried on from Month 37:

- Stage 1: initial assessment from IT Experts and Clinicians;
- Stage 2: overtime evaluation, carried out throughout the project by Clinicians and Participants.

The obtained results will drive and enhance a redesign process (see Figure 6).

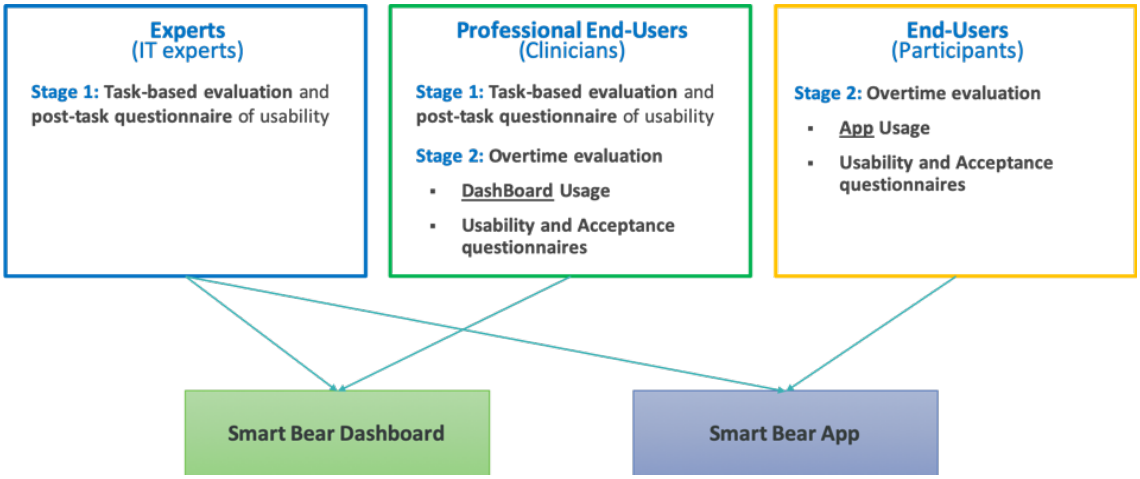

Figure 5. Participant Involvement in the two Evaluation Stages

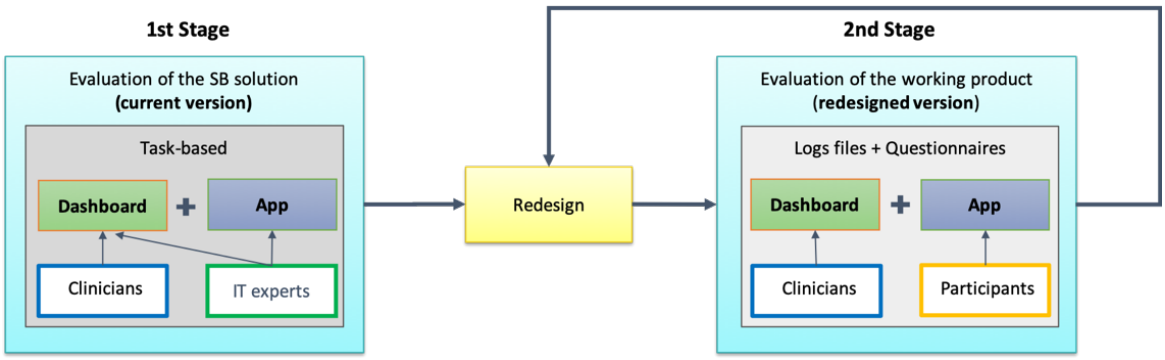

Figure 6. The two Evaluation Stages: a redesign can be carried out according to the obtained results

Three measures will be considered:

- Usability: defined in terms of effectiveness, efficiency, satisfaction according to ISO 9241-11 guidance;
- Technology Acceptance: reception of SB technology and intention to adopt it;
- Usage: when and how the users have interacted with the app/dashboard, percentage of active users, inactive users, drop out.

The different measures will be assessed using standardized tools accessible through the App/Dashboard and performing a statistical analysis of the collected data.

5.2 Stage 1 Procedure

An initial assessment will be obtained by performing an evaluation task scenario on the SB@App (mock-up version)/Dashboard (testing environment).

A minimum number of two clinicians from each SB clinical partner and two IT experts from each Research or Industrial partner will be involved in the evaluation.

The participants (IT experts -> App and Dashboard, Clinicians -> Dashboard), will be invited to carry out a set of specific tasks to investigate how they approach the app/dashboard sections and functionalities. Each participant will be provided with a short description of the tasks to be completed, as well as the app/dashboard manual to be consulted in case of need.

A number of 8 tasks have been defined (details are reported in Section 7.1.2):

- Task on SB App:
  - Task1@App: Diet App: monthly habits questionnaire;
  - Task2@App: Diet App - daily food intake;
  - Task3@App: OTAGO/rehabilitation;
  - Task4@App: medication Intake.
- Tasks on SB dashboard:
  - Task 1@Dashboard: setting of a patient's health habits;
  - Task2@Dashboard: setting of a patient's medications;
  - Task3@Dashboard: RGA questionnaire completion;
  - Task4@Dashboard: creation of a new notebook.

After completing each task, the subjects will fill in the ASQ (After Scenario Questionnaire) (Lewis, 1995) and an open-ended question for the elicitation of functionalities and the deepening of issues (details are reported in Section 7.1.2, Table 20).

After the completion of all the tasks, each participant will fill in the Post Study System Usability Questionnaire (PSSUQ) (Lewis, 1995) (details are reported in Section 7.1.2, Table 21). The questionnaire allows for an assessment of system usefulness, information quality, and interface quality. An additional item was included to keep track of the age of each participant. The questionnaires will be accessible through the App/Dashboard. The scores will be stored in the non-FHIR DB (single score per item, total score per questionnaire).

Moreover, during the accomplishment of the task, the following data will be collected and stored in App and Dashboard log files:

- starting and ending timestamp for each task
  - starting point -> opening the Smart Bear App
  - ending point -> all planned actions have been performed (end condition is different for each task)
- {button clicked + event timestamp} pairs for each task, to track all user actions and possible errors from the beginning to the end of the task and verify that the task is performed with the least number of clicks without navigating unnecessary app sections.

The collected data will be analysed to quantitatively estimate:

- Effectiveness: how many users were able to complete the given set of tasks (% tasks solved, % of users able to complete a given task)
- Efficiency: the effort to complete a task (number of clicks to solve the task, time taken to solve the task).

5.3 Stage 2 Procedure

The second stage consists of an overtime evaluation, carried out throughout the project by professional and non-professional end-users (Clinicians->Dashboard, Recruited Participants->App).

During the entire period will be collected usage data corresponding to the interactions with the App/Dashboard. Periodically, the data will be examined and the clinicians will be invited to compile questionnaires related to Usability, Technology Acceptance, and Satisfaction. Two data gathering periods will be considered.

To take into account the specificity of the context, in accordance with the literature, it was deemed appropriate to assess Usability, relying both on a generic and a context-specific questionnaire. The most commonly used generic tools are UMUX (4 items) and SUS (10 items) (Lewis, 1028).

Since the use of two questionnaires would result in an excessive number of items, it was preferable to choose UMUX since it is shorter, and it is also based on the official definition and standard established by the ISO (Lewis, 2018) (details are reported in Section 7.1.3, Table 23).

As a specific questionnaire, it was chosen to use two versions of MAUQ (mHealth App Usability Questionnaire), respectively for the evaluation of the app by Participants and the dashboard by Clinicians (Zhou et al.2019), (details are reported in Section 7.1.3, Table 23 and Table 24).

Although Satisfaction is included in the ISO definition of Usability and is provided in the UMUX tool, the NPS tool was included (Reichheld, 2003) (details are reported in Section 7.1.3, Table 25), as it is composed of only one item that is perfectly aligned with the KPI25.

Technology Acceptance will be evaluated using the standard tool UTAUT (Unified Theory of Acceptance and Use of Technology) (VenKatesh et al., 2003) (details are reported in Section 7.1.3, Table 26). The model explains the intentions of individuals towards the use of technology and the behaviors they assume when adopting it. UTAUT has also the advantage of unifying in a single model most of the constructs already present and validated in previous research (TAM included).

According to the motivations mentioned above, a variation of KPI-13 and i-KPI-17 is proposed. The amendment proposes to use UMUX and UTAUT instead of SUS and TAM.

Usage will be evaluated from the analysis of log files. The aim is to keep track of the active users, sessions length, and activities. It will be considered:

- a coarse-grained usage metric: number of accesses to the app and the dashboard in a specific period of time;
- a fine-grained usage metric: number of sessions, session duration, frequent activities.

5.4 Ethical Guidelines and Requirements

Data will be collected anonymously to ensure that any participant's activities and scores will not be individually attributable. Evaluation will be done by users within the project who fall under the ethical approval process.

The tools to be used in Stage 1 will be administered in English since the participants involved in the task-based scenario evaluation are all professional users accustomed to using English in their work practices. Instead, the tools to be used in Stage 2 will be translated into the languages of the pilots.

6 References

[1] Airaksinen O., Brox JI., Cedraschi C., Hildebrandt J., Klaber-Moffett J., Kovacs F., Mannion AF., Reis S., Staal JB., Ursin H., Zanoli G.; COST B13 Working Group on Guidelines for Chronic Low Back Pain. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J. (2006) Mar;15 Suppl 2(Suppl 2): S192-300. doi: 10.1007/s00586-006-1072-1. PMID: 16550448; PMCID: PMC3454542.

[2] Airaksinen, Olavi & Brox, Jens & Cedraschi, Christine & Hildebrandt, Jan & Klaber-Moffett, J & Kovacs, Francisco & Mannion, AF & Reis, Shmuel & Staal, J Bart & Ursin, H & Zanoli, Gustavo. (2006a). European Guidelines for the Management of Chronic Non-Specific Low Back Pain Cost B13 working groupEur Spine J200615S192S30010.1007/s00586-006-1072-116550448. European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 15 Suppl 2. S192-300. 10.1007/s00586-006-1072-1.

[3] Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas. 2013 May;75(1):51-61. doi: 10.1016/j.maturitas.2013.02.009. Epub 2013 Mar 22. PMID: 23523272.

[4] Angeli, F., Reboldi, G., Verdecchia, P., 2020. Hypertensive urgencies and emergencies: Misconceptions and pitfalls. Eur J Intern Med, 71, p.15–17.

[5] Anguera JA, Boccanfuso J, Rintoul JL, et al. Video game training enhances cognitive control in older adults. Nature. 2013. doi:10.1038/nature12486

[6] Astarita, A., Covella, M., Vallelonga, F., Cesareo, M., Totaro, S., et al, 2020. Hypertensive emergencies and urgencies in emergency departments: a systematic review and meta-analysis. Journal of Hypertension, 38(7), p.1203–1210.

[7] Bellandi, V., Ceravolo, P., Damiani, E. et al. A methodology to engineering continuous monitoring of intrinsic capacity for elderly people. Complex Intell. Syst. (2022). https://doi.org/10.1007/s40747-022-00775-w

[8] Beller, J., Geyer, S. & Epping, J. Health and study dropout: health aspects differentially predict attrition. BMC Med Res Methodol 22, 31 (2022). https://doi.org/10.1186/s12874-022-01508-w

[9] Bhandari, B., Narasimhan, P., Jayasuriya, R., Vaidya, A., Schutte, A.E., 2022. Effectiveness and Acceptability of a Mobile Phone Text Messaging Intervention to Improve Blood Pressure Control (TEXT4BP) among Patients with Hypertension in Nepal: A Feasibility Randomised Controlled Trial. gh, 17(1), p.13.

[10] Boisgontier M. P., Beets I. A., Duysens J., Nieuwboer A., Krampe R. T., & Swinnen, S. P. (2013). Age-related differences in attentional cost associated with postural dual tasks: increased recruitment of generic cognitive resources in older adults. Neuroscience and biobehavioral reviews, 37(8), 1824–1837. https://doi.org/10.1016/j.neubiorev.2013.07.014

[11] Bragazzi, N.L., Zhong, W., Shu, J., Abu Much, A., Lotan, D., et al, 2021. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. European Journal of Preventive Cardiology, 28(15), p.1682–1690.

[12] Broekhuis, M., van Velsen, L., Peute, L., Halim, M., & Hermens, H. (2021). Conceptualizing Usability for the eHealth Context: Content Analysis of Usability Problems of eHealth Applications. JMIR Formative Research, 5(7), e18198.

[13] da Silva TBL, Dos Santos G, Moreira APB, Ishibashi GA, Verga CER, de Moraes LC, Lessa PP, Cardoso NP, Ordonez TN, Brucki SMD. Cognitive interventions in mature and older adults, benefits for psychological well-being and quality of life: a systematic review study. Dement Neuropsychol. 2021 Oct-Dec;15(4):428-439. doi: 10.1590/1980-57642021dn15-040002. PMID: 35509795; PMCID: PMC9018088.

[14] Ernawati, I., Lubada, E.I., Lusiyani, R., Prasetya, R.A., 2022. Association of adherence measured by self-reported pill count with achieved blood pressure level in hypertension patients: a cross-sectional study. Clin Hypertens, 28(1), p.12.

[15] Fan, B. J., & Wong, R. Y. (2019). Effect of cognitive training on daily function in older people without major neurocognitive disorder: a systematic review. Geriatrics, 4(3), 44.

[16] Garcez D. R., da Silva Almeida G. C., Silva C., de Souza Nascimento T., de Athayde Costa E Silva A., Kleiner A., da Silva Souza G., Yamada E. S., & Callegari, B. (2021). Postural adjustments impairments in elderly people with chronic low back pain. Scientific reports, 11(1), 4783. https://doi.org/10.1038/s41598-021-83837-2

[17] Garner J. Conceptualizing the relations between executive functions and self-regulated learning. J Psychol Interdiscip Appl. 2009;143(4):405-426. doi:10.3200/JRLP.143.4.405-426

[18] Gates NJ, Vernooij RWM, Di Nisio M, Karim S, March E, Martínez G, Rutjes AWS. Computerised cognitive training for preventing dementia in people with mild cognitive impairment. Cochrane Database of Systematic Reviews 2019, Issue 3. Art. No.: CD012279. DOI: 10.1002/14651858.CD012279.pub2. Accessed 28 August 2022.

[19] Gatsios Dimitris & Bamiou, Doris-Eva & Costafreda Sergi & Georga, Eleni & Kourou Konstantina & Fotiadis, Dimitrios & Bibas, Athanasios (Thanos) & Nikitas, Christos & Liston, Matthew & Pavlou, Marousa. (2019). Achieving adherence in home-based rehabilitation with novel human machine interactions that stimulate community-dwelling older adults. 616-619. 10.1145/3316782.3322762.

[20] Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, Lamb SE. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev. 2012 Sep 12;2012(9):CD007146. doi: 10.1002/14651858.CD007146.pub3. PMID: 22972103; PMCID: PMC8095069.

[21] Hajesmaeel-Gohari, S., Khordastan, F., Fatehi, F., Samzadeh, H., & Bahaadinbeigy, K. (2022). The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Medical Informatics and Decision Making, 22(1), 1-9.

[22] Hall CD, Herdman SJ, Whitney SL, et al. Vestibular Rehabilitation for Peripheral Vestibular Hypofunction: An Updated Clinical Practice Guideline From the Academy of Neurologic Physical Therapy of the American Physical Therapy Association. J Neurol Phys Ther. 2022;46(2):118-177. doi:10.1097/NPT.0000000000000382

[23] Hampstead BM, Sathian K, Phillips PA, Amaraneni A, Delaune WR, Stringer AY. Mnemonic strategy training improves memory for object location associations in both healthy elderly and patients with amnestic mild cognitive impairment: A randomized, single-blind study. Neuropsychology. 2012;26(3):385-399. doi:10.1037/a0027545

[24] Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized cognitive training in older adults with mild cognitive impairment or dementia: A systematic review and meta-analysis. Am J Psychiatry. 2017;174(4). doi:10.1176/appi.ajp.2016.16030360

[25] Holmes B, Leggett S, Mooney V, Nichols J, Negri S, Hoeyberghs A. (1996) Comparison of female geriatric lumbar-extension strength: asymptotic versus chronic low back pain patients and their response to active rehabilitation. J Spinal Disord, 9(1):17-22.CDC. National Center for Injury Prevention and Control - Home Page|Injury Center|CDC. Centers for Disease Control and Prevention. https://www.cdc.gov/injury/index.html. Published October 4, 2019. Accessed October 14, 2019.

[26] Hu M, Wu X, Shu X, et al. Effects of computerised cognitive training on cognitive impairment: a meta-analysis. J Neurol. 2019;(0123456789). doi:10.1007/s00415-019-09522-7

[27] Huckans, M., Hutson, L., Twamley, E., Jak, A., Kaye, J., & Storzbach, D. (2013). Efficacy of cognitive rehabilitation therapies for mild cognitive impairment (MCI) in older adults: working toward a theoretical model and evidence-based interventions. Neuropsychology review, 23(1), 63-80.

[28] Jilla, A.M., Johnson, C.E., Danhauer, J.L., Anderson, M., Smith, J.N., et al, 2020. Predictors of Hearing Aid Use in the Advanced Digital Era: An Investigation of Benefit, Satisfaction, and Self-Efficacy. J Am Acad Audiol, 31(2), p.87–95.

[29] Koch C., & Hänsel F. (2019). Non-specific Low Back Pain and Postural Control During Quiet Standing-A Systematic Review. Frontiers in psychology, 10, 586. https://doi.org/10.3389/fpsyg.2019.00586

[30] Leontsinis, I., Katsiki, N., Papademetriou, V., Chrysohoou, C., Kariori, M., et al, 2021. Hypertensive urgencies during the first wave of COVID-19 pandemic in a tertiary hospital setting: A "U" shaped alarming curve. Arch Med Sci. Available at: https://www.archivesofmedicalscience.com/Hypertensive-urgencies-during-the-first-wave-of-COVID-19-pandemic-in-a-tertiary-hospital,141243,0,2.html [Accessed July 29, 2022].

[31] Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. International Journal of Human-Computer Interaction, 7(1), 57-78.

[32] Lewis, James R. "Measuring perceived usability: The CSUQ, SUS, and UMUX." International Journal of Human–Computer Interaction 34.12 (2018): 1148-1156.

[33] Lukitasari, M., Nugroho, D., Rohman, M., Mardhotillah, H., Natasya, D., et al, 2021. An intervention study for impact assessment of health education by empowered community health workers in improving treatment and diet adherence in hypertension. Indian J Community Med, 46(4), p.618.

[34] Ma, L.-Y., He, F., Liu, S., Wang, X.-D., Gao, Y., et al, 2022. The Association Between the Prevalence, Medication Adherence and Control of Hypertension and the Prevalence of Mild Cognitive Impairment in Rural Northern China: A Cross-Sectional Study. PPA, Volume 16, p.493–502.

[35] Mancia, G., Rosei, E.A., Azizi, M., Burnier, M., Clement, D.L., et al, 2018 ESC/ESH Guidelines for the management of arterial hypertension. , p.98.

[36] Marshall L. M., Litwack-Harrison S., Cawthon P. M., Kado D. M., Deyo R. A., Makris U. E., Carlson H. L., Nevitt M. C., & Study of Osteoporotic Fractures (SOF) Research Group (2016). A Prospective Study of Back Pain and Risk of Falls Among Older Community-dwelling Women. The journals of gerontology. Series A, Biological sciences and medical sciences, 71(9), 1177–1183. https://doi.org/10.1093/gerona/glv225

[37] Meyerowitz-Katz, G., Ravi, S., Arnolda, L., Feng, X., Maberly, G., & Astell-Burt, T. (2020). Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. Journal of Medical Internet Research, 22(9), e20283. DOI: 10.2196/20283

[38] Michas, G., Magriplis, E., Panagiotakos, D., Chourdakis, M., Micha, R., et al, 2021. Heart failure in Greece: The Hellenic National Nutrition and Health Survey (HNNHS). Hellenic Journal of Cardiology, 62(4), p.315–317.

[39] Nguyen-Huynh, M.N., Young, J.D., Ovbiagele, B., Alexander, J.G., Alexeeff, S., et al, 2022. Effect of Lifestyle Coaching or Enhanced Pharmacotherapy on Blood Pressure Control Among Black Adults With Persistent Uncontrolled Hypertension: A Cluster Randomized Clinical Trial. JAMA Netw Open, 5(5), p.e2212397.

[40] Peng, Z., Jiang, H., Wang, X., Huang, K., Zuo, Y., Wu, X., ... & Yang, L. (2019). The efficacy of cognitive training for elderly Chinese individuals with mild cognitive impairment. BioMed research international, 2019.

[41] Pothula V. B., Chew F., Lesser T. H., & Sharma A. K. (2004). Falls and vestibular impairment. Clinical otolaryngology and allied sciences, 29(2), 179–182. https://doi.org/10.1111/j.0307-7772.2004.00785.x

[42] Puts, M. T. E., Toubasi, S., Andrew, M. K., Ashe, M. C., Ploeg, J., Atkinson, E., ... & McGilton, K. (2017). Interventions to prevent or reduce the level of frailty in community-dwelling older adults: a scoping review of the literature and international policies. Age and ageing, 46(3), 383-392.

[43] Reichheld, F. F. (2003). The one number you need to grow. Harvard business review, 81(12), 46-55.

[44] Rosano, G.M.C., Seferovic, P., Savarese, G., Spoletini, I., Lopatin, Y., et al, 2022. Impact analysis of heart failure across European countries: an ESC-HFA position paper. ESC Heart Failure, p.ehf2.14076.

[45] Schiltenwolf M., Akbar M., Hug A., Pfüller U., Gantz S., Neubauer E., Flor H., & Wang H. (2014). Evidence of specific cognitive deficits in patients with chronic low back pain under long-term substitution treatment of opioids. Pain physician, 17(1), 9–20.

[46] Schiltenwolf M., Akbar M., Neubauer E., Gantz S., Flor H., Hug A., & Wang H. (2017). The cognitive impact of chronic low back pain: Positive effect of multidisciplinary pain therapy. Scandinavian journal of pain, 17, 273–278. https://doi.org/10.1016/j.sjpain.2017.07.019

[47] Seferović, P.M., Vardas, P., Jankowska, E.A., Maggioni, A.P., Timmis, A., et al, 2021. The HEART FAILURE ASSOCIATION ATLAS : HEART FAILURE EPIDEMIOLOGY AND MANAGEMENT STATISTICS 2019. European J of Heart Fail, 23(6), p.906–914.

[48] Senoo, K., Miki, T., Ohkura, T., Iwakoshi, H., Nishimura, T., et al, 2022. A Smartphone App to Improve Oral Anticoagulation Adherence in Patients With Atrial Fibrillation: Prospective Observational Study. JMIR Mhealth Uhealth, 10(1), p.e30807.

[49] Shanbehzadeh S., Salavati M., Talebian S., Khademi-Kalantari K., & Tavahomi M. (2018). Attention demands of postural control in non-specific chronic low back pain subjects with low and high pain-related anxiety. Experimental brain research, 236(7), 1927–1938. https://doi.org/10.1007/s00221-018-5267-6

[50] Sheppard, J.P., Albasri, A., Gupta, P., Patel, P., Khunti, K., et al, 2021. Measuring adherence to antihypertensive medication using an objective test in older adults attending primary care:

cross-sectional study. J Hum Hypertens. Available at: https://www.nature.com/articles/s41371-021-00646-w [Accessed August 25, 2022].

[51] Simon, S. S., Yokomizo, J. E., & Bottino, C. M. (2012). Cognitive intervention in amnestic Mild Cognitive Impairment: a systematic review. Neuroscience & Biobehavioral Reviews, 36(4), 1163-1178.

[52] Skelton D, Dinan S, Campbell M, Rutherford O. Tailored group exercise (Falls Management Exercise -- FaME) reduces falls in community-dwelling older frequent fallers (an RCT). Age Ageing. 2005 Nov;34(6):636-9. doi: 10.1093/ageing/afi174. PMID: 16267192.

[53] Śliwińska-Kowalska, M., Zaborowski, K., 2017. WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Permanent Hearing Loss and Tinnitus. International Journal of Environmental Research and Public Health, 14(10), p.1139.

[54] Stookey, A. D., & Katzel, L. I. (2020). Home exercise interventions in frail older adults. Current geriatrics reports, 9(3), 163-175.

[55] Suh J. H., Kim H. Jung G. P, Ko J. Y., Ryu J. S. (2019). The effect of lumbar stabilization and walking exercises on chronic low back pain A randomized controlled trial. Medicine (Baltimore), 98(26): e16173.

[56] Tam, H.L., Leung, L.Y.L., Wong, E.M.L., Cheung, K., Chan, A.S.W., 2022. Integration of text messaging interventions into hypertension management among older adults: A systematic review and meta-analysis. Worldviews Ev Based Nurs, 19(1), p.16–27.

[57] Tecca JE. Are post-fitting follow-up visits not hearing aid best practices? Hearing Review. 2018;25(4):12-22.

[58] Tromp, J., Beusekamp, J.C., Ouwerkerk, W., Meer, P., Cleland, J.G.F., et al, 2022. Regional differences in precipitating factors of hospitalization for acute heart failure: insights from the REPORT-HF registry. European J of Heart Fail, 24(4), p.645–652.

[59] Van Daele U., Hagman F., Truijen S., Vorlat P., Van Gheluwe B., & Vaes P. (2010). Decrease in postural sway and trunk stiffness during cognitive dual-task in nonspecific chronic low back pain patients, performance compared to healthy control subjects. Spine, 35(5), 583–589. https://doi.org/10.1097/BRS.0b013e3181b4fe4d

[60] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

[61] Viken H, Reitlo LS, Zisko N, Nauman J, Aspvik NP, Ingebrigtsen JE, Wisløff U, Stensvold D. Predictors of Dropout in Exercise Trials in Older Adults: The Generation 100 Study. Med Sci Sports Exerc. 2019 Jan;51(1):49-55. doi: 10.1249/MSS.0000000000001742. PMID: 30113524Grover S, Mehra A, Chakrabarti S, Avasthi A. Dropout rates and reasons for dropout from treatment among elderly patients with depression. J Geriatr Ment Health 2018;5:121-7

[62] Wan, J., Wu, Y., Ma, Y., Tao, X., Wang, A., 2022. Predictors of poor medication adherence of older people with hypertension. Nursing Open, 9(2), p.1370–1378.

[63] Whitmer, W.M., Howell, P., Akeroyd, M.A., 2014. Proposed norms for the Glasgow hearing aid benefit profile (GHABP) questionnaire. Int J Audiol, 53(5), p.345–351.

[64] Yun, S., & Ryu, S. (2022). The Effects of Cognitive-Based Interventions in Older Adults: A Systematic Review and Meta-Analysis. Iranian Journal of Public Health, 51(1), 1.

[65] Zhang, L., Sun, X., Liao, L., Zhang, S., Zhou, H., et al, 2019. Effectiveness of blood pressure-lowering treatment by the levels of baseline Framingham risk score: A post hoc analysis of the

Systolic Blood Pressure Intervention Trial (SPRINT). J Clin Hypertens (Greenwich), 21(12), p.1813–1820.

[66] ZHOU, Leming, et al. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR mHealth and uHealth, 2019, 7.4: e11500.

7 Appendix

7.1 Usability Evaluation Framework

7.1.1 Summary of the evaluation tools retrieved in the literature

Table 15 summarises the tools retrieved from the literature review. For each tool are specified:

- the acronym;
- the name of the tool;
- the target user;
- the target measure;
- the number of items: for some tools, it is possible to adapt the number and descriptions of items, the table indicates the most commonly used versions according to the literature analysis;
- the specificity for the health domain;
- the publication year;
- the evaluation scale.

Table 15. Tools retrieved in the literature.

Acronym	Name	Intended Users	Measure	Items	Health-Specific	Year	Scale
AdEQUATE	QuestionnAire for Evaluation of QUAlity in TElemedicine system	end-user (no professional)	quality	67	yes	2015	4-Point Likert
ASQ	After Scenario Questionnaire	all	satisfaction	3	no	1995	7-point Likert
CSQ-8	Client Satisfaction Questionnaire	end-user (no professional)	satisfaction	8	no	1979	4-Point Likert
CSUQ	IBM Computer system usability questionnaire	end-user (no professional)	usability	19	no	1993	7-point Likert
Health-ITUES	Health InformationTechnology Usability Evaluation Scale	end-user (no professional)	usability	20	yes	2010	5-point Likert
MARS	Mobile App Rating Scale	health care professional	quality	23	yes	2015	5-point Likert
MAUQ	mHealth App Usability Questionnaire	all	usability	18	yes	2019	7-point Likert
mHealth Satisfaction Questionnaire	mHealth Satisfaction Questionnaire	end-user (no professional)	satisfaction	14	yes	2020	5-point Likert

NASA TLX	NASA Task Load Index	all	workload	6	no	1988	7-point Likert
NAU	Nielsen's Attributes of Usability	all	usability	5	no	1993	7-point Likert
NHE	Nielsen's Heuristic Evaluation	all	usability	10	no	1993	7-point Likert
NPS	Net-Promoter Score (NPS)	all	satisfaction	1	no	2003	10-point
Perceived Ease of Use and Usefulness	Perceived Ease of Use and Usefulness Questionnaire	end-user (no professional)	usability	12	no	1989	7-point Likert
PHUE	Practical Heuristics for Usability Evaluation	all	usability	13	no	1997	7-point Likert
PSSUQ	Post Study System Usability Questionnaire	all	usability	18	no	1992	7-point Likert
PUEU	Perceived Usefulness and Ease of Use	professional user	usability	12	no	1989	7-point Likert
PUTQ	Purdue Usability Testing Questionnaire	all	usability	100	no	1997	7-point Likert
QUIS	Questionnaire of User Interface Satisfaction	all	satisfaction	27	no	1988	10-point
SEQ	Single Ease Question	all	usability	1	no	2009	7-point Likert
SUS	System Usability SCALE	all	usability	10	no	1996	5-point Likert
TAM	Technology Acceptance Model	all	acceptance	12	no	1989	7-point Likert
TUQ	Telehealth Usability Questionnaire	end-user (no professional)	usability	21	yes	2016	7-point Likert
UEQ	User experience questionnaire	end-user (no professional)	usability	26	no	2008	7-point Likert
uMARS	User Mobile App Rating Scale	end-user (no professional)	quality	20	yes	2016	5-point Likert
UMUX	Usability Metric for User Experience	end-user (no professional)	usability	4	no	2010	5 or 7-point Likert
USE	The usefulness, satisfaction, and ease of use	all	usability	30	no	2001	7-point Likert
UTAUT	Unified theory of acceptance and use of technology	all	acceptance	31	no	2003	7-point Likert

7.1.2 Stage 1

In what follows, the tasks defined for the app/dashboard evaluation scenario and the assessment tools are described.
For each task, are specified:

- the start and end time of the task;
- the description provided to the subjects;
- a table with the path of the expected actions to be performed by the subjects.

7.1.2.1 Descriptions of tasks for the app evaluation

Task 1@App.
Start Time: click on the diet app;
End Time: submit the questionnaire on eating habits.
The description provided to the participant: "*Fill in your eating habits. Use the app manual in case of difficulties".*

Table 16. Task 1@App: Diet App - Monthly Habits Questionnaire

Task 1@App: Diet App - Monthly Habits Questionnaire
1. Open the Diet App from the HomePage
2. Access the habit questionnaire from the icon on the top right of the Diet App
3. Fill in the monthly questionnaire about your eating habits
4. Press "SUBMIT"
5. Fill in the ASQ evaluation survey
6. Fill in the open-ended question

Task 2@App.
Start Time: click on the diet app;
End Time: button submit pressed on all the food categories in the DAILY section.
The description provided to the participant: "*Fill in your daily food intake. Use the app manual in case of difficulties* ".

Table 17. Task 2@App: Diet App - Daily food intake

Task 2@App: Diet App - Daily food intake
1. Open the Diet App from the HomePage
2. Fill in your DAILY food intake by tapping on ALL the food categories and fill in how many portions you consumed
3. Fill in the ASQ evaluation survey
4. Fill in the open-ended question

Task 3@App.
Start Time: click on the balance app;
End Time: all video exercises have been watched.

The description provided to the participant: *"Imagine you are in a situation of "no pain, no recent falls, no shortness of breath, level of unsteadiness less than 5. In the balance app, browse through all rehabilitation exercises videos and fill in the associated questionnaires. Use the app manual in case of difficulties"*.

Table 18. Task 3@App: OTAGO/rehabilitation - Monthly Questionnaire Scenario

Task 3@App: OTAGO/rehabilitation - Monthly Questionnaire Scenario
1. Open the Balance App from the HomePage
2. Open the Rehab App
3. Fill in the Questionnaire with the following choices: No Pain, No Shortness of Breath, No Fall, 0-5 Steadyness.
4. Read the Instructions
5. Watch all the rehabilitation exercises videos and fill in the questionnaires between them
6. Fill in the ASQ evaluation survey
7. Fill in the open-ended question

Task 4@App.
Start Time: click on the medication app;
End Time: timeframes confirmed ("ok" button pressed).
The description provided to the participant: "*For all proposed medications, select 9:00 and 21:00 as timeframes for each medication. Use app manual in case of difficulties"*.

Table 19. Task 4@App: Medication Intake

Task 4@App: Medication Intake
1. Open the Medication App from the HomePage
2. For each medication in the list, change the time intervals for taking medication by selecting 9:00 and 21:00
3. Fill in the ASQ evaluation survey
4. Fill in the open-ended question

7.1.2.2 Descriptions of tasks for the dashboard evaluation

Task 1@Dashboard:
Start Time: click on the patient tab;
End Time: save the data.
The description provided to the subjects: "*Show the evaluation of the patient with id=1, change and save his health habits by setting: high salt assumption, current everyday smoker, one pack of cigarettes per month, two falls in the past 12 months, one alcoholic drink per day. Use the manual of the dashboard in case of difficulties. Use the dashboard manual in case of difficulties*".

Table 20. Task 1@Dashboard: Setting of a patient's health habits

Task 1@Dashboard: Setting of a patient's health habits
1. Open the patients' tab from the HomePage
2. Search for the patient with id=1
3. In the evaluation section, show the patients' evaluations
4. In the habits section, change the health habits by setting: high salt assumption, current everyday smoker, one pack of cigarettes per month, two falls in the past 12 months, one alcoholic drink per day
5. Save the inserted data
6. Check if the information has been successfully updated
7. Fill in the ASQ evaluation survey
8. Fill in the open-ended question

Task 2@Dashboard:
Start Time: click on the patient tab;
End Time: save the data.
The description provided to the subjects: "*For the patient with id=1, add the medication Domiphen 0.5 mg every 4 hours for three days, meals oral route and then add the nutritional supplements Vitamin D and Vitamin D derivative. Use the dashboard manual in case of difficulties*".

Table 21. Task 2@Dashboard: Medications Setting

Task 2@Dashboard: Medications Setting
1. Open the patients' tab from the HomePage
2. Search for patient with id= 1
3. In the evaluation section, show the patient's evaluations
4. In the medications section, add the following medications: Domiphen 0.5 mg every 4 hours for three days, meals oral route
5. In the nutritional supplement section, add the following supplements: vitamin D and vitamin D derivative
6. Save the inserted data
7. Check if the information has been successfully updated
8. Fill in the ASQ evaluation survey
9. Fill in the open-ended question

Task 3@Dashboard:
Start Time: click on the patient tab;
End Time: save the data.
The description provided to the subjects: "*Compile the RGA questionnaire for the patient with id=1, by simulating the responses provided by the patient. Use the dashboard manual in case of difficulties".*

Table 22. Task 3@Dashboard: Questionnaire Completion

Task 3@Dashboard: Questionnaire Completion
1. Open the patients' tab from the HomePage
2. Search for the patient with id=1
3. In the evaluation section, show the patient's evaluations
4. In the questionnaires section, fill in the RGA questionnaire by simulating the responses provided by the patient
5. Save the inserted data
6. Check if the information has been successfully updated
7. Fill in the ASQ evaluation survey
8. Fill in the open-ended question

Task 4@Dashboard:
Start Time: click on the notebook tab;
End Time: check the data.
The description provided to the subjects: "*Create a new notebook from the notebook tab. Use the dashboard manual in case of difficulties*".

Table 23. Task 4@Dashboard: Notebook (AI)

Task 4@Dashboard: Notebook (AI)
1. Open the notebook tab from the HomePage
2. Create a new notebook
3. Check if the information has been successfully updated
4. Fill in the ASQ evaluation survey
5. Fill in the open-ended question

7.1.2.3 Descriptions of the evaluation tools

After Scenario Questionnaire (ASQ)

The items are reported in Table 24.
The score is evaluated as the average of the three responses.
The following open-ended question has been included: "*Describe your overall experience with the app highlighting strengths or weaknesses and any suggestions for improvement*".

Table 24. After Scenario Questionnaire

ASQ (After Scenario Questionnaire)	1 (Strongly Disagree)	2	3	4	5	6	7(Strongly Agree)	NA
Overall, I am satisfied with the ease of completing the task in this scenario.								
Overall, I am satisfied with the amount of time it took to complete the task in this scenario.								
Overall, I am satisfied with the support information (on-line help, messages, documentation) when completing the task.								

Post Study System Usability Questionnaire Questionnaire (PSSUQ)

The items are reported in Table 25. The sub-scales provide a more detailed breakdown of different factors affecting the website, software, system or product. The following open-ended question has been included: "*Please write your suggestions to improve the app/dashboard".*
From the answers, it is possible to evaluate:

- Overall score: the average scores of questions 1 to 16;
- System Usefulness (SYSUSE): the average scores of questions 1 to 6;
- Information Quality (INFOQUAL): the average scores of questions 7 to 12;
- Interface Quality (INTERQUAL): the average scores of questions 13 to 15.

*Table 25. Post Study System Usability Questionnaire. * not to show to the subjects*

PSSUQ			1 (Strongly Disagree)	2	3	4	5	6	7 (Strongly Agree)	NA
System Usefulness*	1	Overall, I am satisfied with how easy it is to use this system								
	2	It was simple to use this system								
	3	I was able to complete the tasks and scenarios quickly using this system								
	4	I felt comfortable using this system								
	5	It was easy to learn to use this system								
	6	I believe I could become productive quickly using this system								

Information Quality*	7	The system gave error messages that clearly told me how to fix problems								
	8	Whenever I made a mistake using the system, I could recover easily and quickly								
	9	The information (such as online help, on-screen messages, and other documentation) provided with this system was clear								
	10	It was easy to find the information I needed								
	11	The information was effective in helping me complete the tasks & scenarios								
	12	The organization of information on the system screens was clear								
Interface Quality*	13	The interface of this system was pleasant								
	14	I liked using the interface of this system								
	15	This system has all the functions and capabilities I expect it to have								
	16	Overall, I am satisfied with this system								

7.1.3 Stage 2

7.1.3.1 Descriptions of the evaluation tools

Usability Metric for User Experience (UMUX) Questionnaire

The items are reported in Table 26. The score is obtained by:

- scoring odd items as [user score - 1];
- scoring even items as [7 - user score];
- adding up these differences and dividing the sum by 24 (the highest possible score);
- multiplying the quotient by 100;
- averaging the results across users.

Table 26. Usability Metric for User Experience Questionnaire

UMUX	1 (Strongly Disagree)	2	3	4	5	6	7 (Strongly Agree)	NA
The SB dashboard/ app capabilities meet my requirements.								
Using SB dashboard/ app is a frustrating experience								
The SB dashboard/ app is easy to use.								
I have to spend too much time correcting things with SB dashboard/ app								

mHealth App Usability Questionnaire (MAUQ)

Two versions for the evaluation of the app by participants and the dashboard by clinicians are reported respectively in Table 27 and Table 28. The score is obtained by determining the average of the responses to all statements.

*Table 27. mHealth App Usability Questionnaire for Clinician. * not to show to the subjects*

MAUQ		1 (disagree)	2	3	4	5	6	7 (agree)
Ease of use and satisfaction *	The dashboard was easy to use.							
	It was easy for me to learn to use the dashboard.							
	The navigation was consistent when moving between screens.							
	The interface of the dashboard allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app.							
System information arrangement *	Whenever I made a mistake using the dashboard, I could recover easily and quickly.							

	I like the interface of the dashboard	.						
	The information in the dashboard was well organized, so I could easily find the information I needed.							
	The dashboard adequately acknowledged and provided information to let me know the progress of my action.							
	I feel comfortable using this dashboard in social settings.							
	The amount of time involved in using this dashboard has been fitting for me.							
	I would use this dashboard again.							
	Overall, I am satisfied with this dashboard.							
Usefulness *	The dashboard would be useful for my health care practice.							
	The dashboard improved my access to delivering health care services.							
	The dashboard helped me manage my patients' health effectively							
	This dashboard has all the functions and capabilities I expected it to have.							
	I could use the dashboard even when the Internet connection was poor or not available.							
	This mHealth dashboard provides an acceptable way to deliver healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment.							

*Table 28. mHealth App Usability Questionnaire for Participants: * not to show to the subjects*

MAUQ		**1 (disagree)**	**2**	**3**	**4**	**5**	**6**	**7 (agree)**
Ease of use and satisfaction *	The app was easy to use.							
	It was easy for me to learn to use the app.							
	The navigation was consistent when moving between screens.							
	The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app.							
System information arrangement *	Whenever I made a mistake using the app, I could recover easily and quickly.							
	I like the interface of the app							
	The information in the app was well organized, so I could easily find the information I needed.							

	The app adequately acknowledged and provided information to let me know the progress of my action.							
	I feel comfortable using this app in social settings.							
	The amount of time involved in using this app has been fitting for me.							
	I would use this app again.							
	Overall, I am satisfied with this app.							
Usefulness *	The app would be useful for my health and well-being.							
	The app improved my access to healthcare services.							
	The app helped me manage my health effectively.							
	This app has all the functions and capabilities I expected it to have.							
	I could use the app even when the Internet connection was poor or not available.							
	This mHealth app provides an acceptable way to receive healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment.							

Net Promoter Scale (NPS)

The item of the questionnaire is reported in Table 29. The score is obtained as the difference between the percentage of Promoters and the percentage of Detractors, where:

- Promoters are the respondents who scored the system a 9 or 10;
- Passives are the respondents who have answered a 7 or 8;
- Detractors are the respondents who scored the system in the range of 0-6

Table 29. Net Promoter Scale

How likely are you to recommend SMART BEAR platform to a friend or colleague?”	0	1	2	3	4	5	6	7	8	9	10

UTAUT Unified theory of acceptance and use of technology

Figure 7 and Table 30 describe the model for exploring factors influencing user intentions to adopt a technology and an example of items to estimate the UTAUT. The model includes four core variables (Performance expectancy, Effort expectancy, Social influence, Facilitating conditions) and four

moderating variables that have significant impacts on core variables (gender, age, voluntariness, experience). The items for the evaluation of the SB app and dashboard are in the process of being defined.

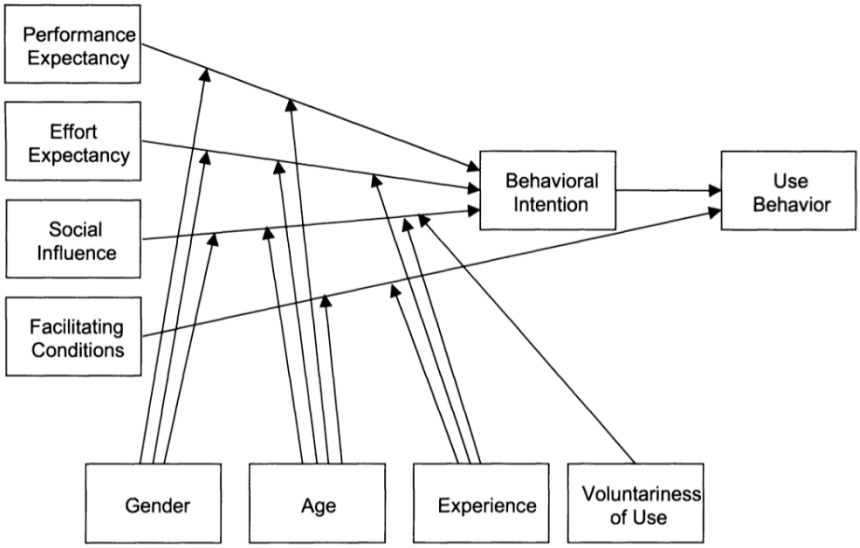

Figure 7. UTAUT model (Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly)

Table 30. UTAUT Items

Construct	Item
Performance expectancy	PE1: I would find the system useful in my job.
	PE2: Using the system enables me to accomplish tasks more quickly.
	PE3: Using the system increases my productivity.
	PE3: If I use the system, I will increase my chances of getting a raise.
Effort expectancy	EE1: My interaction with the system would be clear and understandable.
	EE2: It would be easy for me to become skillful at using the system.
	EE3: I would find the system easy to use.
	EE4: Learning to operate the system is easy for me.
Attitude toward using technology	ATUT1: Using the system is a bad/good idea.
	ATUT2: The system makes work more interesting.
	ATUT3: Working with the system is fun.
	ATUT4: I like working with the system.
Social influence	SI1: People who influence my behavior think that I should use the system.
	SI2: People who are important to me think that I should use the system.
	SI3: The senior management of this business has been helpful in the use of the system.
	SI4: In general, the organization has supported the use of the system.
Facilitating conditions	FC1: I have the resources necessary to use the system.
	FC2: I have the knowledge necessary to use the system.
	FC3: The system is not compatible with other systems I use.

	FC4: A specific person (or group) is available for assistance with system difficulties.
Self-efficacy	I could complete a job or task using the system …
	SE1: If there was no one around to tell me what to do as I go.
	SE2: If I could call someone for help if I got stuck.
	SE3: If I had a lot of time to complete the job for which the software was provided.
	SE4: If I had just the built-in help facility for assistance.
Anxiety	ANX1: I feel apprehensive about using the system.
	ANX2: It scares me to think that I could lose a lot of information using the system by hitting the wrong key.
	ANX3: I hesitate to use the system for fear of making mistakes I cannot correct.
	ANX4: The system is somewhat intimidating to me.
Behavioral intention to use the system	BI1: I intend to use the system in the next <n> months.
	BI2: I predict I would use the system in the next <n> months.
	BI3: I plan to use the system in the next <n> months.