A Flexible Toolkit for Evaluating Person-Centred Digital Health and Wellness at Scale

Marilyn McGee-Lennon, Matt-Mouley Bouamrane, Eleanor Grieve, Catherine A. O'Donnell, Siobhan O'Connor, Ruth Agbakoba, Alison M. Devlin, Sarah Barry, Annemieke Bikker, Tracy Finch and Frances S. Mair

Abstract The Delivering Assisted Living Lifestyles at Scale (dallas) program was a large-scale, nationwide deployment of digital health and wellbeing products and services in the UK. Telehealth, telecare, mobile apps, personal health records, and assisted living technology were implemented by four large multi-stakeholder consortia and a multidimensional evaluation was carried out across the lifecycle from examining co-design and redesign of services through to rolling out services via statutory, private and consumer routes. A flexible toolkit of descriptive, process and

M. McGee-Lennon (✉) · M.-M. Bouamrane (✉)
Computer and Information Sciences, University of Strathclyde, Glasgow, UK
e-mail: marilyn.lennon@strath.ac.uk

M.-M. Bouamrane
e-mail: mattmouley.bouamrane@strath.ac.uk

E. Grieve (✉) · C.A. O'Donnell · S. O'Connor (✉) · R. Agbakoba (✉) · A.M. Devlin (✉) ·
S. Barry (✉) · A. Bikker (✉) · F.S. Mair (✉)
Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK
e-mail: Eleanor.Grieve@glasgow.ac.uk

S. O'Connor
e-mail: s.oconnor.1@research.gla.ac.uk

R. Agbakoba
e-mail: r.agbakoba.1@research.ac.uk

A.M. Devlin
e-mail: alison.devlin@glasgow.ac.uk

S. Barry
e-mail: Sarah.barry@glasgow.ac.uk

A. Bikker
e-mail: Annemieke.Bikker@glasgow.ac.uk

F.S. Mair
e-mail: Frances.Mair@glasgow.ac.uk

T. Finch
Institute of Health and Society, Newcastle University, Newcastle, UK

outcome measures was developed and iteratively refined throughout the program. This approach enabled a longitudinal mixed-methods evaluation, underpinned by a robust social theory of implementation called 'Normalization Process Theory'. There remains uncertainty about the best approaches to real world digital health evaluation. This program provided a unique opportunity to develop the knowledge base and toolkit of qualitative and quantitative methods necessary to evaluate person-centered digital health technologies deployed at scale.

Keywords Health informatics · eHealth · Digital health · Telemedicine · Implementation · Evaluation

1   Introduction

Population demographics are changing with growing numbers of older adults who have complex health and social care needs and a concomitant rise in the prevalence of chronic illness and multi-morbidity (i.e. having two or more chronic illnesses) [1]. As a result, health and social care services are struggling to cope with the increasing care burden. To address this, health and care systems are shifting towards promoting and supporting the management of long-term conditions in the community and supporting older adults to live independently, by encouraging people to become proactive stakeholders in managing their own health and well-being through 'self-care' [2, 3].

To support this 'self-care' agenda, a wide range of person-centered digital health platforms and devices have emerged, including: telehealth and telecare systems [4], personal health records [5], various mobile 'apps' and wearable technologies [6, 7] and assisted living [8]. Many of these technologies have been deployed in a limited context however and firm and generalizable evidence of their benefits is still lacking.

Randomized Controlled Trials (RCTs) are currently viewed as the gold standard in assessing the effectiveness of health interventions and have therefore also been advocated to evaluate digital health interventions [9]. However, there is increasing recognition that these more controlled evaluation methods are not necessarily the most practical or appropriate in the context of consumer and patient digital health (eHealth) implementations [10–12]. These methodologies do not allow us to understand implementation issues in other populations, other geographies or other social, political or financial contexts. They also focus on summative end-point evaluations of stable products rather than understanding how these technologies and services are actually designed, implemented and rolled out and iteratively and formatively evaluating and improving them based on feedback from actual use in context. Alternative types of evaluation methods are required to capture complex, real world digital health 'intervention'. Mixed methods, participatory and theoretically grounded approaches are increasingly being advocated and applied to the field of 'eHealth' [13, 14].

This paper addresses this gap in the literature by describing:

1. A flexible toolkit used to evaluate an evolving large scale, national digital health deployment program;
2. How we used a robust theoretical framework, Normalization Process Theory (NPT) to interpret and understand implementation processes; and
3. How the evaluation had to adapt to meet changing requirements throughout this dynamic program.

2 Background

To address the challenges of understanding how to best scale and routinize digital health, Innovate UK funded a large scale £37 million three-year program in the UK called 'Delivering Assisted Living Lifestyles at Scale' (dallas).[1] The dallas program funded four large multi-stakeholder consortia called 'Living It Up', 'Year Zero', 'More Independent', and 'i-Focus' [15–18]. These were jointly led by a variety of health and social care services, private industry, the voluntary sector, local government agencies and in some cases, supported by academic research partners.

The dallas program aimed to:

(i) Support and evaluate the design, development and implementation of technologies and services that enable citizens to take greater control of their health and wellness and
(ii) Encourage a large-scale consumer market for person-centered technologies by driving social and service innovation.

The program specifically focused on promoting preventive care, self-care and independent living through 6 key values called the '6 Cs': '*Control, Choice, Community, Connectedness, Contribution and Collaboration*'. Each dallas community developed and deployed a broad range of technologies and services from traditional telehealth and telecare systems, personal health records, mobile health apps, assisted-living devices and online digital health portals and activities to raise awareness of self-care and increase digital access and inclusion. These were targeted at a variety of people from patients and consumers, from children to older adults, and in both urban and rural regions of the UK [15–18].

---

[1]https://connect.innovateuk.org/web/dallas.

3 Evaluation Framework

The breadth of technologies and range of services to be deployed in the dallas program were not defined at the outset of the evaluation [15]. Furthermore, the proposed digital platforms earmarked for development as part of the program were intended to cross many organizational and cultural boundaries (health and social care for example) as well as being targeted at multiple geographies and populations [15–18]. From the outset our overall evaluation approach had three aspects as illustrated in Table 1 (descriptive, process and outcome) with a fixed underpinning theoretical framework, Normalization Process Theory (NPT) [19, 20] to help us capture and organize the implementation data. It rapidly became necessary to adopt an agile and participatory approach to designing the *outcome* evaluation methods and toolkit (a suite of evaluation tools and approaches to assess outcomes). This included active involvement and engagement with the communities at events and iterative reflective cycles to understand what was being rolled out, to whom, and how this could be best evaluated (what the evaluation criteria should be for each product and service).

This participatory and reflective process enabled us to develop a practical and robust evaluation approach that met the evolving needs of the program. It meant that our toolkit and approach could be flexible and responsive to external contexts so it could be adapted as the socioeconomic and political landscape rapidly and radically evolved throughout the duration of the program we were evaluating. To ensure we critically and reflexively monitored and adapted the evaluation framework throughout the program, weekly research meetings, monthly project management meetings and regular consultations with the four dallas consortia and program funder were undertaken. An independent advisory board was established which included experts from a range of backgrounds (academic, health service, industry and voluntary sector) as well as a nominated individual from the funding body. This group met annually and provided considerable added value to the evaluation cycles through their collective expertise.

Table 1 Dallas evaluation framework

Descriptive	*Project documentation*—quarterly technical reports, observational logs, local evaluation data *Ethnographic observations*—dallas leads meetings, co-design workshops
Process	*Interviews*—baseline, midline and endpoint e-HIT interviews with a cross section of members from each dallas group *Focus groups*—with patients, carers, consumers and health professionals *Project documentation*—quarterly technical reports, observational logs *Ethnographic observations*—dallas meetings, dallas dissemination events *Surveys*—NoMAD surveys with health professionals
Outcome	*Minimum Data Set (MDS)*—date of birth, gender, postcode *Surveys*—core, bespoke and contingent valuation surveys *Project documentation*—recruitment reports

3.1 Study Design

The products and services being rolled out by the four dallas consortia were diverse, heterogeneous and evolved significantly throughout the duration of the program. For example some services were digital (digital personal health records) and others not (awareness raising initiatives). Some products were aimed at statutory markets (e.g. a health professional might prescribe a fall detector or a heart monitor for example) and some were consumer-based wellness or lifestyle products which one could buy off the shelf, or download from an app store (for example an app to support carers of people with dementia). Each product and service was also targeted at different populations and users ranging from older adults to new mothers to people with specific long term conditions (for example Heart Failure). The resulting evaluation therefore was largely an implementation evaluation (because there was no baseline and no single stable intervention to be studied) but one which also provided insight into how to capture outcome level data at scale in the wild for consumer based digital health programs.

3.2 Multidimensional Evaluation Aims

The evaluation aims were to understand the potential impacts of the program on:

 (i) Individuals (including end users, their friends and family and formal carers)
(ii) Systems (such as healthcare, housing social care).

   This required us to explore tools and instruments for gathering data on user experiences (whether the services were usable and acceptable for example), potential outcomes (whether quality of life changed because of the service), and experiences of key implementers designing, developing and rolling out the various services and products (for example barriers and facilitators to uptake and roll out). The evaluation had to be multidimensional and therefore we created a framework composed of the three aforementioned dimensions (see Table 1).

   Firstly, a dynamic descriptive evaluation allowed us to examine 'Who', 'Why', 'Where' and 'When' in relation to the range of digital health interventions developed and also how these evolved over time. Secondly, a longitudinal process evaluation focused on the 'How'; providing insights into the implementation journey and the way it was shaped by complex mechanisms and external factors. Thirdly, an outcomes evaluation aimed to capture the 'What', using quantitative data in order to capture objective measures of the reach (who was accessing and using products and services) and effectiveness (actual and perceived benefits and views of using a service or product) of the dallas program.

3.3 Theoretical Underpinning

The evaluation was underpinned by a robust sociological theory of implementation processes called Normalization Process Theory (NPT) [19, 20]. NPT—in the context of change and innovation within complex organisations—provides a useful prism through which one can observe, identify, analyze and explain the variety of factors which affect how individual, collective and organizational 'agency' impact new work practices [19, 20]. NPT has four core generative mechanisms: *coherence, cognitive participation, collective action* and *reflexive monitoring*. This provided a framework through which the dallas implementation processes (captured via observations, interviews and surveys) could be analyzed and conceptualized in a robust and coherent manner (Table 2). NPT was chosen as much of its early development was undertaken in the context of studies of implementation of eHealth services and trials and it has proven useful as a tool across a broad range of contexts.

4 Evaluation Toolkit

This section presents an overview of the tools, methods or instruments used to collect data for each of the three evaluation approaches described above (descriptive, outcomes, implementation). In each, we describe what data was being collected and how; what this data allowed us to evaluate and then present some insight into the benefits and limitations of each tool or method. It is not intended that this toolkit should be used in its entirety for each digital health evaluation; *rather that people can see the value of developing a flexible toolkit based on their own aims and objectives and practical limitations*.

Table 2 The four constructs of normalization process theory (NPT)

Coherence	Cognitive participation	Collective action	Reflexive monitoring
The work of understanding and "making sense" of new practices or ways of working	The work of engaging individuals and groups to 'buy into' and adapt the organization to new practices	The work of implementing new practices and providing the necessary resources and training to operationalize these	The work of evaluating and monitoring new practices and appraising whether they are worth sustaining or not

4.1 Descriptive Evaluation: Tracking the Evolution of Products and Services

@@While the different communities adopted varying levels of co-design (co-operation and collaboration with people who might use the service) in their approaches, all involved solution exploration and user-centered design approaches from the outset [16–18] in order to design and redesign services that people needed and wanted. This meant that the exact specification of a product remained unclear while services were being designed. While this is the very purpose of co-design—i.e. getting to 'the heart' of what people really want from a product and potentially increasing eventual buy-in and take-up later on—it also posed considerable evaluation challenges.

Defining appropriate measures of success is much more straightforward with a stable system or defined service rather than ones which are evolving or effectively 'work-in-progress'. Learning about the benefits and challenges of this co-design phase itself seemed therefore an essential part of this evaluation. To accommodate this continuously evolving landscape we documented product and service development activities: how they changed over time, identifying 'critical transition points' and reasons and motivations for adaptation. This product development 'journey' is a useful research output of its own accord. Creating a set of evaluation criteria for such a diverse and dynamic set of services is a considerable challenge. Hence, the set of tools we selected needed to include both generic measures—i.e. some which capture general usability and wellness—as well as more sensitive tools where appropriate for context-specific interventions: for example, self-efficacy scales for self-management apps.

4.2 Counting Users and Measuring Reach

One of the first challenges for deployment evaluations is to agree on an objective and practically useful definition of a participant (involved in the program in some way) as well as levels of engagement (active and meaningful involvement). In conventional trials it is often straightforward to count a participant once they are assigned to a service or experimental 'condition' and "consent" to study participation. In consumer based trials however, users can be actively recruited to a service or choose to use a product or service at a time of their choosing and of their own accord (e.g. installing software or downloading an app). It is thus harder to measure and track when a person becomes an active participant and have one single baseline for evaluation. Furthermore, there is a lack of clarity in a consumer based trial regarding what the term 'engaged user' means. If a person downloads an app or registers for a service they should certainly be 'counted' as being a participant, but at what point do they become considered active and engaged. This could be based on length of time registered (someone who stays registered for the duration of a

weight-loss program for example), or on patterns of usage data (someone that logs in at least once a week for example).

For this evaluation an individual was considered '*reached*' if they engaged with a dallas product or service (regardless of whether they were currently using the product or service). This might have been active participation at a workshop or event where the products were being designed or marketed or passive engagement in terms of TV, Radio, or newspaper advertisements or posters or leaflets in a community setting. An individual was considered a '*member*' of the program when they could be counted and described in terms of their basic demographics (age, gender, socioeconomic status) which allowed us to evaluate and report on the reach of each product or service. Finally, an individual was considered a 'user' if there was direct evidence of them signing up to use a product or service and there was evidence of use (varying from logging in once, to setting up a personal profile, to having a full set of logged usage data for a product). Different categories of user might be affected by the program in different ways and therefore it is important to decide what the different outcomes might be for such a wide range of participation.

4.3   Process Evaluation

A primary aim of this evaluation was to capture key stages of the various products and services as they were designed, developed, piloted and implemented and to document and chart this in a way that would allow us to identify real barriers and facilitators for implementation of digital health at scale. Understanding barriers, facilitators and key lessons learned across dallas was a key evaluation goal from the outset as such work would provide valuable "best practice" for future large-scale technology deployments. In order to capture the full breadth of barriers and facilitators across the stakeholders and organizations in the program we collected extensive implementation data as shown in Table 3.

Using this data set and the underlying NPT framework to understand and categorize the data we were able to explore many complex barriers and facilitators to the uptake of digital health. Some of these have already been reported [18] and key lessons are being disseminated [21]. An important point was the consistent use of our theoretical framework, NPT, to underpin data collection and analysis across the program.

There is currently great interest in the factors which contribute to, or impede, the implementation of complex interventions at scale and in real world settings. Our series of in-depth, e-HIT led interviews (a tool that has NPT as its theoretical underpinning) with a purposive sample of key stakeholders from each of the four 'dallas' communities allowed us to track key implementation facilitators and or barriers in a consistent fashion. Qualitative data analysis from these longitudinal interviews showed us that key stakeholders recognized the role that the program had played in enabling them to operationalize aspects of service re-design to deliver more personalized, digital tools and services for citizens and service users which

Table 3 The implementation data collected as part of the dallas implementation evaluation

E-HIT (e-Health Implementation Toolkit) interviews (N = 46):
with key informants, stakeholders and implementers (Baseline (n = 17); Mid-point (n = 21); and End-point (n = 10) of 3 year program) to explore implementation issues
Key Implementer Interviews (N = 52) and Champion Interviews (N = 23):
to capture expectations and experiences of stakeholders involved in championing digital self-care in their community
Quarterly Monitoring reports (N = 48):
submitted to Innovate UK and analysed by us to track development of services, products and activities and identify barriers and facilitators
Quarterly Evaluation reports (N = 43):
to track services, products and activities and identify barriers and facilitators to deployment
Recruitment reports and Observation Logs (N = 11) and Meeting minutes and observation field notes:
to capture recruitment numbers, strategies and blockers and facilitators, to capture on the ground service roll out blockers and facilitators, to track development of services, products and activities and identify barriers and facilitators
NoMAD surveys (N = 153):
to explore implementation issues with health professionals, undertaken in two services

included sharing of knowledge to help sustainability at a wider program level. However, it also permitted active emergence of knowledge about implementation challenges that were noted across the life of this real-world, large scale digital health deployment.

In addition to the longitudinal set of key implementer interview we also conducted surveys with health professionals to assess "readiness" to adopt new technologies within the healthcare environment. These were undertaken using a specialized normalization survey called NoMAD (normalization of complex interventions-measure development). The NoMAD team worked with the dallas consortia to adapt the survey questions to each service and identify appropriate staff for inclusion in the respective surveys. NoMAD questions were derived from NPT, again highlighting the importance given to consistent use of theory in this evaluation, something that has been deemed important in digital health evaluations [13].

Participants (health professionals involved in rolling out services) were asked to respond to items on a 5-point Likert scale from 'strongly agree' to 'strongly disagree'. The data for each site was analyzed descriptively, using response frequencies and grouped according to professional roles. The NoMAD tool provided useful insights into health professionals' perceptions during the implementation process for certain initiatives where it was used and the findings from it resonated with the large volume of qualitative data collected in parallel as part of the wider dallas evaluation. Responses from NoMAD surveys were generally positive about engagement; but there was scope for improving understanding about the potential benefits of each initiative. The need for adequate provision of training and resources was also highlighted.

The NoMAD tool helped demonstrate that by listening to implementers, key learning can be revealed which can help allow the smooth integration of the initiative. Longitudinal qualitative data was analyzed using a framework underpinned by NPT. However, data was not "shoehorned" to fit this framework as we allowed for identification of themes that fell outside the framework to maximize knowledge generation about implementation issues and ensure that the theory did not constrain learning.

4.4   Measuring Outcomes

With a variety of digital health products all with different potential end users and different functionality it was difficult to define a clear set of primary outcomes to measure across the whole program. Due to the evolving nature of the program itself it became clear that measuring health related outcomes would be challenging as:

(i)   although the program was of 3 years duration many initial offerings had to be reconsidered and it was only in the final phase of the program that the services for evaluation were more clearly defined and operational

(ii)   it was difficult to define outcomes when products and services were meant to positively affect different people (end users, friends and family and carers) in different ways (increase independence, enable them to self-manage, improve their sense of control);

(iii) it was hard to attribute any measured benefits to the program when there was no "control" group for comparison.

For these reasons a suite of outcome measures were developed and this outcomes based toolkit could be used by communities to collect baseline data and potentially in the future to collect follow-up data once they had stable products with users signed up to use them. Those involved in the development and deployment of consumer digital health are not necessarily familiar with the evaluation methods required to demonstrate impact and outcomes. Companies on the one hand might want to measure metrics such as downloads and the number of kits sold whereas the statutory sector might be more interested in service use and sustainability of services. Funders and commissioners however still rely heavily on outcome-based reporting and therefore it was important that we developed methods by which people could collect data on who was using their product; the perceived usability and usefulness of the products/services and whether or not using such as service or product had an effect on their lives. At the outset this toolkit was envisaged as consisting of traditional outcome measures such as EQ-5D (to measure health related quality of life). However, by the end of the program it had evolved to consist of the following items:

- A 'Rate this Product/Service' item
- Questions on how the person found out about a product/service

- a 'Minimum Dataset' (gender, date of birth, post code, ethnicity)
- *to capture 'reach'*
- questions on computer, internet and smartphone usage
- questions on Perceived usability and Perceived Usefulness *based on the validated Technology Acceptance* Model [22]
- Perceived influence on lifestyle factors
- Health directed behaviors [23].

4.5 Measuring Usability and User Experience

To measure usability and user experience ten items from the system usability and perceived usefulness elements on the Technology Acceptance Model survey were used. It was important for us to provide a measure of both of these elements in case products were either seen to be really useful but were not usable or acceptable or the product was highly usable but not something that would be readily taken up by the intended user group. Due to the evolving nature of the program there was limited time to collect actual user experience data via the survey because products and services were only being piloted and still refined and developed and we did not have direct access to users during the program, who were recruited via the consortia themselves.

In addition to the survey items it was important for us to conduct more qualitative evaluations inside the communities with some of the users of their deployed services and products. We therefore conducted 8 focus groups with 59 service users in order to understand what positive and negative experiences people had in relation to using the various digital health services being rolled out across dallas. This was only a sample and—although it shed important light on some of the user engagement issues that can be experienced—should not be generalized to all digital health products or roll outs.

4.6 Measuring Quality of Life and the Value of 'Wellness'

Many of the services provided within the dallas program were centered around the consumer 'wellness' market rather than 'health service delivery'. Some standard measures—such as health related quality of life measures [24]—were perceived as inadequate to the 'consumer well-being' market context. Companies rolling out wellness apps for example did not want to ask traditional questionnaire items that focus heavily on statutory care and health status as opposed to general wellbeing.

Evaluating health and wellness products and services with a range of benefits and which are not purely focused on direct health gains will mean that other wellness based measures need to be developed and validated and that further

exploration of methodologies to assess "cost-effectiveness" or the "value" of new digital health services is urgently required. To address this we included a measure of 'lifestyle wellness' in the dallas evaluation (*the dallas 6 Cs: contribution, choice, control, collaboration, community and connectedness*). This constituted both a risk and an opportunity. Although not a validated questionnaire item, it was essential to measure these aspects of the program given the novel and holistic nature of dallas.

Health economists have also highlighted the inherent difficulty in assigning personal values to enhancing 'well-being'. To address this, a population contingent valuation survey was developed as part of the dallas evaluation in order to understand what value the public might put on these lifestyle attributes [25]. This in turn enabled us to identify the "public's" willingness to pay (WTP) for digital interventions to promote health and wellbeing by directly valuing the attributes of the 6C's which were integral to the dallas products and services.

We recognize the strengths and weaknesses of using traditional (validated and accepted) instruments versus novel (no validation, harder to make comparisons across studies) measures but fully believe that new measures of wellness at scale are required for the future of digitally enabled health and wellness programs. Quality Adjusted Life Years (QALYs) [26] are unlikely to be the most appropriate outcome for assessing the benefits achieved from deployment of digital consumer wellness products. There is a need to value a wider range of benefits of consumer wellness products which go beyond health gains by using more consumer facing approaches like discrete choice experiments and contingent valuation studies in order to better evaluate their potential impact on both individuals and health and care systems.

5   Conclusion and Future Work

There is a pressing need to develop flexible but coherent evaluation frameworks to evaluate complex, consumer-centric digital health implementations. Demonstrating the potential impact of health and wellness technologies is crucial as we move into a time where care models are shifting to embrace digitally enabled self-care as routine practice [27, 28]. This paper presented an evaluation framework adopted during the dallas program that allowed us to study digital health consumer technologies at scale in 'real world' setting in a useful way. Our resulting evaluation framework enabled us to provide:

   (i)   a rich descriptive evaluation of the community and service journeys;
  (ii)   learning about the barriers and facilitators to implementation and processes for change;
 (iii)   data on the reach of dallas products and services and
  (iv)   a better understanding of the readiness for digital self-care in the UK. While:
   (v)   The use of a theoretical framework, NPT, enabled us to transform descriptive qualitative data into more explanatory results and make clear recommendations.

References

1. Barnett, K., Mercer, S.W., Norbury, M., Watt, G., Wyke, S., Guthrie, B.: Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet **380**, 37–43 (2012)
2. Davis, M., Freeman, M., Kaye, J., Vuckovic, N., Buckley, D.: A systematic review of clinician and staff views on the acceptability of incorporating remote monitoring technology into primary care. Telemedicine e-Health **20**(5), 428–438 (2014)
3. Robinson, L., Gibson, G., Kingston, A., Newton, L., Pritchard, G., Finch, T., Brittain, K.: Assistive technologies in caring for the oldest old: a review of current practice and future directions. Aging Health **9**(4), 365–375 (2013)
4. Finkelstein, S., Speedie, S., Potthoff, S.: Home telehealth improves clinical outcomes at lower cost for home healthcare. Telemedicine e-Health **12**(2), 128–136 (2006)
5. Archer, N., Fevrier-Thomas, U., Lokker, C., McKibbon, K.A., Straus, S.E.: Personal health records: a scoping review. J. Am. Med. Inf. Assoc. **18**, 515–522 (2011)
6. Mosa, A.S.M., Yoo, I., Sheets, L.: A systematic review of healthcare applications for smartphones. BMC Med. Inform. Decis. Mak. **12**, 1–31 (2012)
7. Orwat, C., Graefe, A., Faulwasser, T.: Towards pervasive computing in health care—a literature review. BMC Med. Inform. Decis. Mak. **8**, 26 (2008)
8. Linskell, J., Bouamrane, M.-M.: Assisted-living spaces for end-users with complex needs: a proposed implementation and delivery model. Health Inform. J. **18**(3), 159–170 (2012)
9. Steventon, A., Bardsley, M., Billings, J., Dixon, J., et al.: Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ **344**, e3874 (2012)
10. Dansky, K.H., Thompson, D., Sanner, T.: A framework for evaluating eHealth research. Eval. Program Plann. **29**(4), 397–404 (2006)
11. Lilford, R., Foster, J., Pringle, M.: Evaluating eHealth: how to make evaluation more methodologically robust. PLoS Med. **6**(11), e1000186 (2009)
12. Baker, T.B., Gustafson, D.H., Shah, D.: How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J. Med. Internet Res. **16**(2), e36 (2014)
13. Greenhalgh, T., Russell, J.: Why do evaluations of eHealth Programs Fail? An alternative set of guiding principles. PLoS Med. **7**, e1000360 (2010)
14. Finch, T., May, C., Mair, F., Mort, M., Gask, L.: Integrating service development with evaluation in telehealthcare: an ethnographic study. BMJ **327**, 1205–1209 (2003)
15. McGee-Lennon, M., Bouamrane, M-M., Barry, S. et al.: Evaluating the delivery of assisted living lifestyles at scale (dallas). In: Proceedings of HCI 2012—People and Computers XXVI, The 26th BCS Conference on Human Computer Interaction, DIILT'12, Birmingham, UK, 12–14 Sept 2012
16. Agbakoba, R., McGee-Lennon, M., Bouamrane, M.-M., Watson, N., Mair, F.: Implementation factors affecting the large-scale deployment of digital health and well-being technologies: a qualitative study of the initial phases of the 'Living-It-Up' program. Health Inform. J. 1460458215594651 (2015)
17. Agbakoba, R., McGee-Lennon, M., Bouamrane, M.-M., Watson, N., Mair, F.: Implementing a national scottish digital health and wellbeing service at scale: a qualitative study of stakeholders' views. Stud. Health Technol. Inform. **216**, 487–491 (2015)
18. Devlin, A.M., McGee-Lennon, M., O'Donnell, C.A., Bouamrane, M.M., et al.: Delivering digital health and wellbeing at scale: lessons learned during the implementation of the dallas program in the UK. J. Am. Med. Inform. Assoc. **23**(1), 48–59 (2015)
19. May, C., Mair, F., Finch, T., MacFarlane, A., et al.: Development of a theory of implementation and integration: normalization process theory. Implement Sci. **4**, 9p (2009)
20. May, C., Finch, T.: Implementing, embedding, and integrating practices: an outline of normalization process theory. Sociology **43**(3), 535–554 (2009)

21. McGee-Lennon, M., Bouamrane, M-M., Devlin, A., O'Connor, S., et al:. Readiness for digital health: findings from dallas, a National Digital Health Innovation Program. Under Peer-Rev. (2016)
22. Mathieson, K.: Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. Inf. Syst. Res. **2**, 173–191 (1992)
23. Osborne, R.H., Elsworth, G.R., Whitfield, K.: The health education impact questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ. Couns. **66**, 192–201 (2007)
24. EuroQol Group: EuroQol: a new facility for the measurement of health related quality of life. Health Policy **16**, 199–208 (1990)
25. McIntosh, E., Clarke, P., Frew, E., et al.: Applied Methods of Cost-Benefit Analysis in Health Care, Vol 4, 119–138. Oxford, UK: Oxford University Press (2010)
26. Räsänen, P., Roine, E., Sintonen, H., Semberg-Konttinen, V., Ryynänen, O.P., Roine, R.: Use of quality-adjusted life years for the estimation of effectiveness of health care: a systematic literature review. Int. J. Technol. Assess. Health Care **22**, 235–241 (2006)
27. Lupton, D.: Critical perspectives on digital health technologies. Sociol. Compass **8**, 1344–1359 (2014)
28. Sands, D.Z., Wald, J.S.: Transforming health care delivery through consumer engagement, health data transparency, and patient-generated health information. Yearb. Med. Inf. **9**(1), 170–176 (2014)