Opportunities and barriers for adoption of a decision-support tool for Alzheimer's Disease

MAURA BELLIO, Centre for Medical Image Computing, Department of Computer Science, UCL; UCL Interaction Centre

DOMINIC FURNISS, UCL Interaction Centre

NEIL P. OXTOBY, Centre for Medical Image Computing, Department of Computer Science, UCL

SARA GARBARINO, Inria, Universite Cote D'Azur, France

NICHOLAS C. FIRTH, Centre for Medical Image Computing, Department of Computer Science, UCL

ANNEMIE RIBBENS, icometrix, Belgium

DANIEL C. ALEXANDER, Centre for Medical Image Computing, Department of Computer Science, UCL

ANN BLANDFORD, UCL Interaction Centre

Clinical decision-support tools (DSTs) represent a valuable resource in healthcare. However, lack of Human Factors considerations and early design research has often limited their successful adoption. To complement previous technically focused work, we studied adoption opportunities of a future DST built on a predictive model of Alzheimer's Disease (AD) progression. Our aim is two-fold: exploring adoption opportunities for DSTs in AD clinical care, and testing a novel combination of methods to support this process. We focused on understanding current clinical needs and practices, and the potential for such a tool to be integrated into the setting, prior to its development. Our user-centred approach was based on field observations and semi-structured interviews, analysed through workflow analysis, user profiles, and a design-reality gap model. The first two are common practice, whilst the latter provided added value in highlighting specific adoption needs. We identified the likely early adopters of the tool as being both psychiatrists and neurologists based in research-oriented clinical settings. We defined ten key requirements for the translation and adoption of DSTs for AD around IT, user, and contextual factors. Future works can use and build on these requirements to stand a greater chance to get adopted in the clinical setting.

CCS Concepts: • **Human-centered computing** → **HCI design and evaluation methods**; • **Applied computing** → **Health informatics**.

Additional Key Words and Phrases: diffusion of innovation, technology adoption, user-centred design, healthcare, design-reality gap

1 INTRODUCTION

Clinical decision-support tools (DSTs) have the potential to aid specialists dealing with complex conditions, characterised by a quantity of data that is not always "knowledge rich" [1]. Medical data users can be overwhelmed by the volume and heterogeneity of inputs they have to handle, which inhibits their decision making. These tools are normally powered by innovative algorithms,

allowing users to see patterns and trends from big datasets that would not be detectable otherwise, and supporting important medical decisions [30].

A major hurdle is facilitating DST adoption in clinical practice. Many innovations fail when moved to the clinical context [4, 27, 60]. As noted by others, failure can be caused by unsuccessful integration with the existing workflow [49], social context, and trust in the system [60], which can all be summarised as a lack of Human Factors considerations [60]. This is particularly true for DSTs, as they are developed and deployed to support specific tasks for specific medical conditions.

In our work, we aim to explore opportunities and barriers to the design and development of a future novel DST for Alzheimer's Disease (AD). AD is a chronic condition that would benefit from DSTs, given the burden it places on society and the challenges in interpreting the volume and heterogeneity of data over a prolonged time frame. In collaboration with the EuroPOND project [17], which is developing predictive models to forecast AD progression over time, we studied how an innovative DST could be conceptualised and integrated with current clinical practice. For this project, we focus on a specific model for disease progression and staging called an Event-Based Model (EBM) [19, 62]. This data-driven algorithm combines information from different data sources into a longitudinal and individual picture of disease development, offering clear opportunities for clinical utility (supporting clinicians in synthesising data from multiple sources, helping with assessing staging and progression, and stratification of patients).

Defining the potential for translation considering the scope, end users, and context prior to any technical implementations will provide key guidelines and recommendations to inform future system development, not only saving time and resources, but also enhancing the chances of the resulting tool being ultimately used. We adopted a user-centred perspective to study current needs and practices around AD clinical management and decision-making. By using a novel combination of exploratory methods (workflow analysis, user profiles, and design-reality gap assessment), we framed the problem around barriers and opportunities in translating predictive algorithms for clinical use in AD care. We identified the likely early adopters amongst AD healthcare professionals, and we proposed requirements to adopt this technology in the field. This initial set of recommendations will inform the development of a DST that will integrate the EBM in the care workflow. To our knowledge, this is the first time that opportunities for adoption are explicitly studied in the development of a DST for AD. These requirements can be used and extended in further work on similar tools for AD.

2 RELATED WORK

2.1 Clinical Decision Support Tools

Clinical DSTs are software which, based on a given input, produce outputs that clinicians can use to support their decision-making [7, 53, 61]. "DSTs" is an umbrella term to identify a broad range of tools, each with different underlying rules, output, and purpose. Horsky et al. [29] describe six main types of DSTs, including notifications, guidelines, and clinical pathways. Berlin et al. [7] propose a taxonomy to classify DSTs, the categories being: context, knowledge, data source, decision support, information delivery, and workflow. Advantages of DSTs are not only related to the knowledge brought to the clinical routine. Sutton et al. [53] have outlined in their review a number of benefits that go beyond diagnostic and decision support, touching on aspects such as workflow and documentation improvement, patient safety, reduction of human error and cognitive workload, cost containment, and administrative automation. Whilst some DSTs are knowledge-based, meaning they are built on rules deriving from clinical evidence, the vast majority of tools are non-knowledge based, making use of data mining techniques such as Artificial Intelligence (AI), Machine Learning (ML), or pattern recognition [53].

2.2 AI systems for healthcare and predictive models

Large volumes of biomedical data are being generated every day. This phenomenon is supported by more advanced computing technologies, such as imaging techniques, storage space, or simply by the increased availability of tracking sensors [6, 38]. The higher demand, however, does not always come with improved knowledge, as "data can be a great asset to healthcare organisations, but they have to be first transformed into information" [35]. In fact, ways to turn big data into "smart data" [22] are needed. To this end, data analytics and Artificial Intelligence (AI) can uncover patterns in noisy and high dimensional datasets, to produce actionable knowledge. These techniques are not new to computational researchers, who have deployed them for years in business or finance, but the health sector is a relative latecomer. The causes lie in healthcare's peculiar complexities, such as heterogeneity and variable quality of data, high-dimensionality, specialised domain knowledge, and more stringent privacy and security requirements, to mention a few [44]. AI has to adapt to a whole new set of challenges, but in doing so it promises to deliver unprecedented value to healthcare: to foster precision medicine; to generate tailored diagnostics and treatment plans; to predict readmission to hospital; or to monitor progression for disease subtypes [38]. One particular application of AI in medicine is the development of *predictive models*, which "can use patient-specific information to predict the outcome of interest and to thereby support clinical decision-making" [6]. While clinicians still use their intuition and expertise to integrate information from patient history and measured data into predictions, data scientists have tackled the problem in a variety of ways. Basic statistics include *regression* [25], where clinical markers at one time are used to predict values at a later point. Recent advances in AI and ML have moved the discipline towards increasingly automated processes, especially in assessing associations, classifications, and relationships amongst data [1], based on supervised and unsupervised ML techniques. Supervised learning techniques learn mappings from input data to outcome from a large set of labelled data; unsupervised learning aims to identify patterns from unlabelled data. Disease Progression Models (DPMs) are unsupervised learning techniques that aim to learn patterns of disease progression, i.e. temporal series of changes, from data sets with no a-priori labelling of patient stage. Some DPMs use discrete staging, such as the EBM [19, 62]; others aim to capture continuous temporal evolution, such as continuous models [14, 42], or spatio-temporal models [50].

2.3 Alzheimer's Disease and the EBM

DPMs offer potential benefit in understanding and managing dementia. Alzheimer's Disease, the most common cause of dementia, is a global emergency: it affects over 46 million people in the world [51], causing a substantial socioeconomic burden. Moreover, causes and treatments are still not well understood. What is clear is that therapeutic interventions are most likely to be effective at early stages, before symptom onset and irreversible neurodegeneration have occurred. Various clinical indicators, called *biomarkers*, are considered reliable detectors for these early stages, and they come in different forms, depending on the source of data: medical imaging, cognitive and clinical tests, analysis of biological fluids, and many others. DPMs can integrate information from large cross-sectional data sets of biomarkers into quantifiable longitudinal pictures of disease development, at both a population and a personalised level.

In this work, as part of the collaboration with the EuroPOND Project, we consider the EBM. Its simplicity and robustness lends itself to translation, as evidenced by its growing usage in a number of disease areas, such as Huntington Disease [59], Multiple Sclerosis [16], but particularly to our case in AD [19, 62]. The EBM conceives disease progression as a sequence of discrete, irreversible events. Every new event is created when a biomarker value moves outside the standard range, with no option of returning back (an assumption based on the degenerative nature of the disease).

Therefore, the course of the disease is represented by the order in which biomarkers leave the standard range of values, thus predicting the probable future course of the disease over time. Although the EBM is only one amongst many disease progression models emerging over recent years, translating this into a useful and usable tool would pave the way for more sophisticated approaches to follow through subsequently.

2.4   Human factors considerations for DST design

In recent decades, almost all DSTs have failed when moved to clinical practice [15, 31, 33, 60, 61]. Looking at possible causes, it becomes clear that these tools need transparency and explainability even more than those in other sectors, because of personal data handling, ethics, and legal requirements [28]. A common pitfall, though, is that DST development often lacks Human Factors considerations [60]. The importance of Human Factors is illustrated in the work of Yang et al [61], who studied the existing decision process around a heart-pump implant, to identify opportunities for the integration of a new DST. Lack of considerations such as poor fit within the workflow, context, and attention to clinical needs [61] open a gap between design and reality. According to Heeks [27], the severity of this gap can be assessed along seven dimensions that are representative of the users in the context and can predict the likelihood of the system failing or succeeding. The bigger the gap, the greater the change needed to successfully adopt the innovation and, consequently, the chances of failure [27]. This framework reinforces the importance of realigning "work as imagined" and "work as done" [8], a necessary effort that innovators should consider in the "design thinking" phase to deliver a system that is truly fit for purpose [8]. Involving the intended users from the early stages of the design process promotes efficiency, effectiveness and safety [37, 40, 43], according to user-centred design principles (ISO 9241-210:2010 [13]). This approach is fundamental, as the design of a single DST is uniquely linked to the medical condition and tasks that the tool will support, as well as greatly dependent on the end users' beliefs, context, and current practice. A user-centred design approach was adopted from the very early stages of development in ADappt [57], where AD clinical specialists and patients' feedback were iteratively included in the design and implementation of an AD diagnostic tool. Whilst every application differs, some general guidance on how to best design such tools have been reported in reviews [29, 34, 47, 53], the main areas being attributes and quality of the technology (e.g. information quality, information delivery), usability (user control, design conventions, trust), and structure (context, workflow integration, and interoperability).

2.5   Technology adoption and barriers

While implications for design can be considered an established contribution in HCI projects, user-centred design would benefit from considering technology adoption in the case of emerging technologies, to support design thinking and identify key research questions when translating from research to practice. Lindley et al. [41] suggest an early evaluation of implications for adoption to prevent failure, referring to research-based speculations on the future potential adoption of a novel technology. This concept is often linked to the rhetoric of a "proximate future" [5], particularly affecting technology innovations that could only exist in a very controlled environment, without a vision of translating and integrating them into the real world. Thinking of adoption from the first stages of DST development will set the path for a long-term vision. However, this approach requires a change in perspective, the focus being on the potential of the product and its contextual fit with the environment. However, as Lindley et al. note [41], there is little guidance on how to capture and present this concept in HCI projects.

Frameworks on the analysis of technology adoption already exist, such as the Technology Acceptance Model (TAM) [12], the Theory of Planned Behaviour (TPB) [3], or Innovation Diffusion

Theory (IDT) [52]. However, these have been developed for already implemented technologies, and are not suitable for our design concept where a safety-critical diagnostic tool has not yet been designed in detail. We need reasonable confidence that the innovation would be a valuable asset to the intended users, and methods to identify possible routes to adoption of a future design concept by clinicians or researchers involved in AD care.

From a practical perspective, Gagnon et al. [21] propose a set of factors (barriers and facilitators) to the implementation of a new system in the medical setting. These factors are derived from case studies on IT adoption by clinical professionals and are overall divided into four main areas: *factors related to IT*, *individual and professional factors*, *human environment*, and *organisational environment*. This work demonstrates that the most important facilitators in IT adoption are represented by users' perceptions of technology potential and the *context* of implementation. Another important requirement is the impact of a novel technology on the *workflow* [23]. It is necessary to fit the novel system into an existing workflow, without disrupting users' beliefs, but promoting safety, data quality, and work efficiency.

Technology adoption leads to the importance of *diffusion of innovation*. Rogers [52] defines it as *"the process in which an innovation is communicated through channels over time among the members of a social system"*. Diffusion of innovation is influenced by multiple factors (context [18], leaders and champions [24], or network [24]). First and foremost, however, the innovation process is guided and influenced by people, where a particular key role is covered by "early adopters". Lettl et al. [40] describe how this group of users can unlock and disseminate the take up of radical innovations. Early adopters are often skilled professionals [48], embedded in interdisciplinary contexts, with facilitated access to specialised resources, and whose needs anticipate those of their peers, making them motivated and open to innovation [40]. Therefore, we need to identify this user group and their context in relation to the tool's purposes. They can facilitate the introduction of the target innovation in the workflow, collaborate in defining elements to better adapt the technology to a wider context, and diffuse the innovation to their peers.

2.6 Contribution of this work

In this project, we studied opportunities and barriers to the adoption of a DST specifically intended for AD clinical practice. As part of the EuroPOND project, we focused on a tool built on the EBM, with the long term goal of using this research to inform the design and development of a DST that is usable, fit for purpose, and adoptable. Technical papers on the EBM anticipated its potential clinical utility as providing a new understanding of the disease to help clinicians put measurements into perspective, providing more accurate prognosis and a staging system that may guide treatment, care choices, and communication with patients. However, previous work has been technically focused, neglecting the need to clarify whether these tools are realistic and, if so, how we should best package and present relevant information to exploit its potential.

This work addresses an unmet need in the burgeoning field of DSTs for healthcare [15, 31, 33, 60, 61], and to our knowledge this is the first time that opportunities for adoption are explored for a DST to be used by AD specialists. This study makes two main contributions. (1) Providing a better understanding of adoption opportunities and barriers for DSTs intended for AD clinical use; (2) proposing and testing a novel combination of HCI methods that can guide adoption research in the early stages of a tool's development. The methods used (workflow analysis, user profiles, and design-reality gap assessment [27]) were aimed to identify the likely early adopters amongst AD clinical specialists, and how the proposed technology might enhance their work. Like a few other studies [43, 61], we engaged with clinical participants (in our case, AD specialists) from an early stage of the development process, prior to any prototype development. This should shape the future design and translation process into healthcare to fit clinical needs and context, and

promote explainability and trust. The outcome from this work is relevant to engineers and clinical researchers working on DSTs for AD care, to use and build on, whilst setting up an adoption and clinical integration strategy from the early stages of the development process.

3 METHODS

3.1 Approach

Methods selected in this work aim to understand users' current needs and practices and identify possible early adopters and contexts of use. This is done for an interactive system that has not yet been developed. We focussed on the following three questions: (Q1) How is work currently done in AD clinical management?; (Q2) What are the characteristics and needs of front-line clinicians in AD?; (Q3) What are the perceived challenges and potential for adoption of the tool? These findings enabled us to define the key implications for the EBM to express its potential in AD clinical decision-making, and the early adopters and context that are best placed to initially benefit from the EBM.

Workflow analysis. Successful integration of innovative technologies in a real world setting requires adaptation of an existing workflow [60, 63]. Thus, we decided to conduct workflow analysis to understand phases in the assessment of an AD patient [45], that will also support design considerations [55]. We generalised the sequence of tasks using a box-and-arrows representation.

User profiles. We described users' characteristics and needs through the definition of profiles that synthesise ethnographic research on a cohort of similar users [39].

Heeks' design-reality gap model (DRGm). Heeks' model is a semi-quantitative assessment of the possible failure or success of a novel technology, along seven dimensions that are representative of a complex socio-technical system [26, 27]. It has the advantage of being applied pre- and post-hoc. We adopted this method to assess how barriers and needs within the considered settings might influence the adoption of a possible system, before any design commitments.

3.2 Data collection

We conducted *field observations* and *semi-structured interviews*. To identify potential early adopters, the settings for observations and participant recruitment covered different options, with an eye on people orientated towards research and innovation. Observations were conducted at multi-disciplinary discussion team (MDT) meetings, venues where a team of clinicians with different expertise meets to discuss complex clinical cases, including interpretation of imaging scans and treatment planning [32]. Semi-structured interviews involved neurologists and old age psychiatrists (recognised as front line specialists in the tasks for which the models are most likely to be beneficial), from hospitals and memory clinics. Ethical clearance to conduct the studies refers to Project ID UCLIC/1213/015 (Principal investigator Ann Blandford).

Field observations. A total of 6 MDT observation sessions were conducted at the National Hospital for Neurology and Neurosurgery, Queen Square (London). This setting was chosen to familiarise with the terminology and the issues discussed by clinicians, data mentioned, features of the disease, and the workflow. Approximately 10 to 15 medical professionals attended each meeting, including a lead neuro-radiologist, neurologists, training students and others from related disciplines. At each meeting, 13 patients were discussed on average (min 6, max 22), within approximately one hour. Participants gave verbal informed consent to being observed. It was not possible to audio or video record the meeting, since patients' names were read out, so data was recorded in the form of hand-written notes. Relevant information to record were identified through preliminary meetings

with doctors, and iteratively adjusted as more meetings were attended. Field notes included data regarding: general workflow, questions asked by the audience, use of data and images, terminology and concepts describing the disease, and decision-making process. Finally, a neurologist, who was not participating in the study, helped to clarify ambiguous clinical terminology or statements that were made during the observations.

Semi-structured interviews. Six clinicians specialising in AD were interviewed: 5 neurologists and 1 psychiatrist. Participants belonged to first referral ($n=2$) and secondary and tertiary referral ($n=4$) centres, located in the London area ($n=5$) and Belgium ($n=1$). For privacy purposes and due to the small number of participants, affiliations are not mentioned here.

Participants were contacted through MDT meetings and word-of-mouth. Prior to the interview, participants were provided with information sheets, informed consent, and a background questionnaire, to track the variety of the sample. Interviews had a dual purpose: to clarify insights from field observations, and explore the perceived value of predictive models. Questions related to the first point addressed both individual clinical practice and MDT sessions (inspired by the work of Lanceley et al. [36]). Regarding the second point, participants were prompted to reflect on the EBM's potential through a set of visual stimuli based on EBM visualisations used by the developers. These included six different examples of output screens (two examples are shown in Fig. 1) with incremental complexity of information provided. The fixed content for all screens was the sequence of stages generated by the EBM and the classification of the patient at a certain point of the scale. This represented the minimum information, as reported in the left example of Fig. 1. Additional information was introduced gradually in the subsequent visuals, by reporting the list of markers becoming abnormal at each stage, the probability curve of a patient being classified at a particular stage, and the positional variance for each marker. The example on the right of Fig. 1 represents the case with the highest complexity. These were used to stimulate discussion with participants, who were asked to comment on the visual stimuli and reflect on their understanding and applicability to the clinical scenario. Interviews were piloted with two neuro-radiologists and adjusted based on feedback. All interviews lasted between 40 and 70 minutes and all but two were recorded and transcribed. For non-recorded interviews, detailed notes were collected and checked with the interviewees. Participants were offered a free trial of the system once it becomes available.
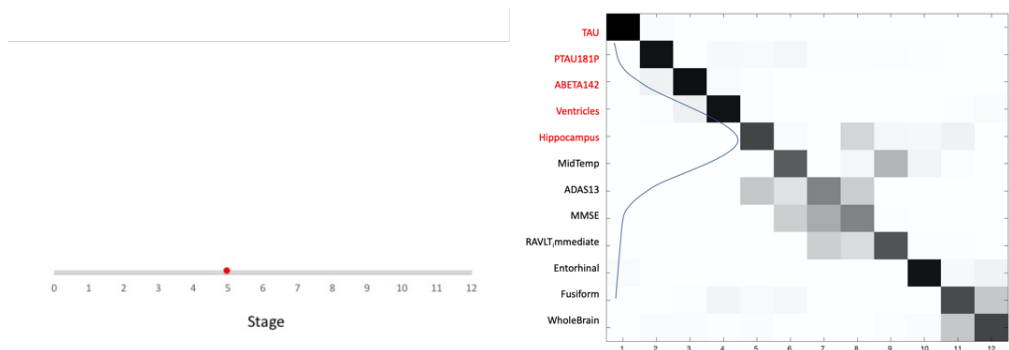

Fig. 1. Two examples of the six visual stimuli used in the semi-structured interviews (credits: *icometrix* early implementation).

3.3 Analysis

The methods listed and their order of application are the result of an exploratory process, whilst addressing our main questions. Firstly, the notes collected during field observations were used to generate a written report. This report grouped the observed information around three different themes: main topics discussed, general patients discussion pattern, and clinical terminology. The content from the report informed the script for the semi-structured interview, which was structured in three main parts: (1) MDT meetings and the process of patients discussion; (2) Participant's individual practice; (3) Introduction to EBM and discussion around its potential and limitations. We first conducted an inductive thematic analysis on field notes and transcribed interviews (software used: NVivo 12): the content was analysed by generating a set of 20 codes that were then grouped into 7 main themes. This first round of analysis revealed correspondence of themes with the factors reported in Gagnon's work [21]. As Gagnon's factors represent a more consolidated framework, they were then used as themes to run a second top-down analysis. All Gagnon's sub-factors were covered by the analysis, except for 5 third-level factors out of 35 and 16 fourth-level factors out of 32. This is due to the fact that interviews were not conducted based on this framework, and also because these factors aim to be generalisable, making it unlikely that a single case study would cover all of them. Findings from the top-down thematic analysis were used to apply the methods described above, which aimed to address our main questions, in chronological order: *workflow analysis* contributed to the understanding of how work is currently done in AD management, *user profiles* gave insights into the characteristics and needs of front-line clinicians in AD, and *Heeks' design-reality gap model (DRGm)* assessed the perceived challenges and potential for adoption of the EBM technology. To provide more strength to the DRGm assessment, we paired Heeks' dimensions with relevant Gagnon's factors, as reported in Table 4.

4 RESULTS

Results follow the set of questions outlined in section 3.1, where each question is addressed by a particular method. Illustrative quotations are identified by participant number (*P1, P2, ...*) and clinical context (*care-oriented: C-O or research and care: R+C*) as described below, to maintain anonymity of participants.

4.1 Current clinical practice and needs

Q1: How is work currently done in AD clinical management? We performed a standard *workflow analysis* of AD current clinical practice, and in particular of AD patients' diagnosis and prognosis, information the EBM is designed to support. We produced a workflow for MDTs and one for individual clinical practice; their combination is represented in Fig. 2. The process starts from availability of adequate data to support a diagnostic decision. Participants reported relying primarily on clinical history and symptoms, followed by cognitive assessment, and then biological and imaging analyses. The first barrier in the process (Fig. 2, point 1) was identified in the challenge of gathering the breadth and quality of data to run the EBM. Typical issues on accessing and interpreting imaging scans were: long waiting lists, delays in obtaining scans, frequent lack of radiologists' reports or interpretation inaccuracies, and low quality images. However, clinicians belonging to research-orientated clinics reported having access to higher quality data, and reports from specialised radiologists. In the subsequent step, a provisional diagnosis is made, combining various cross-sectional data, and led by experience. However, (point 2, Fig. 2) in cases of critical diagnostic uncertainty, further investigations or MDT should be considered. The MDT, on the right side of the map, amplifies the individual expertise and is recognised by all clinicians as a support to clinical uncertainty. The case discussion includes a precise request to the radiologist and to
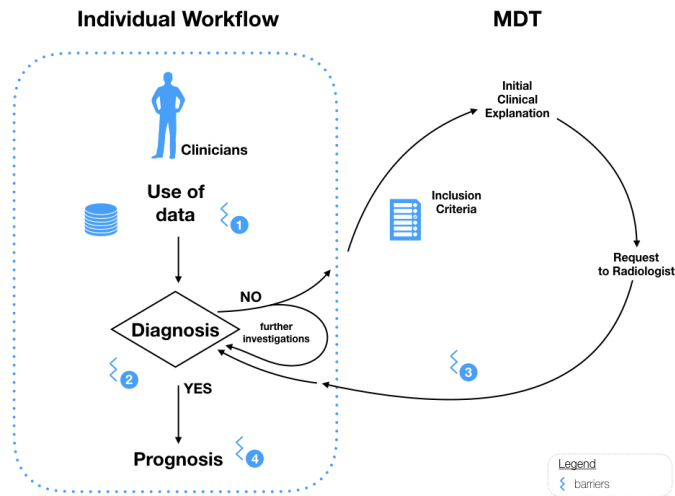
Fig. 2. Combination of workflow's representations: individual and MDT. Barriers are represented by numbers 1 - 4. (1) Availability and sufficient quality of data; (2) Diagnostic uncertainty; (3) Lack of outcome from MDT; (4) Uncertainty in prognosis.

peers, usually to interpret or compare medical imaging data, and advice on further clinical action. The efficiency of MDTs (point 3, Fig. 2) can be disrupted by lack of data clarity or difficulties in converging to a conclusion, leading again into the loop of the diagnosis' decision node. When a provisional diagnosis is fulfilled, a prognosis might be attempted (point 4, Fig. 2). Clinicians confirmed that they reason about prognosis, depending on their confidence in the current diagnosis. However, some interviewees focus more on how they communicate this information to patients, whilst others' interest is in understanding the progression of the disease to more severe stages. The first ones tend not to be too precise in reporting the prognosis, to avoid giving misleading information. The progression of the disease largely depends on its past course and individual differences. One neurologist declared:

> "Generally [the principle is that] the speed [at which it] has started will continue. Sometimes is quite difficult to define." (P3, C-O)

Other clinicians, more interested in early detection of the disease and generally belonging to research-orientated clinics, placed more importance on their ability to detect critical conditions early on, and to monitor patients at risk until a diagnosis is made.

> "The other fundamental point is progression. Understanding what percentage of MCI (mild cognitive impairments) will convert. Because if there is a high probability that a patient will convert from MCI or AD, we would keep them." (P6, R+C)

While MDT workflow is built on consistent responses from all participants, some discrepancies emerged with respect to the individual workflow. Therefore we investigated the variations in context for the clinical specialisations, to clarify how this can influence their needs and, consequently, the technology translation.

Q2: What are the characteristics and needs of front-line clinicians in AD?. To answer this question, we developed three user profiles, characterised by a set of parameters and based on

Table 1. Comparison of three user profiles

Description	Profile 1 *Neurologist, C-O*	Profile 2 *Neurologist, R+C*	Profile 3 *Psychiatrist, R+C*
Resources	Clinical data Cognitive tests (basic) Imaging (low quality, delays)	Clinical data Cognitive tests (complete) Imaging (high quality, delays)	Clinical data Cognitive tests (complete) Imaging (high quality, delays)
Context characteristics	Limited time and resources Patients: all stages of disease If uncertainty: sent to later referral	Available time and resources Patients: uncertain and rare cases Discharged at diagnosis	Available time and resources Patients: at early stages Discharged at diagnosis
Needs	Diagnosis Patient management	Diagnosis Understand disease progression Understand unclear and early cases	Diagnosis Understand disease progression Understand unclear and early cases

outcomes from Q1, observations, and interview data. Table 1 compares and contrasts representative parameters (needs, context, and resources) for the main contexts identified in Q1: *Care-Oriented* (C-O), and *Research+Care* (R+C). A C-O context includes general memory clinics, where patients are referred directly from the primary healthcare provider. Here patients present at very different stages, the visit time is limited, and there is restricted or no access to specialised data collection procedures; routine data collection mainly involves clinical and cognitive tests, while imaging is of variable quality and may or may not be informative for diagnosis and prognosis. R+C refers to specialised memory clinics, research centres, and tertiary care in general. These settings are more frequently involved in uncertain cases and early detection of the disease, until diagnostic criteria are met and the patient is discharged. Access to specialised data is facilitated and systematically recorded.

As per Table 1, firstly we explored whether, considering the same clinical professional role, but in two different contexts, differences emerged around needs, procedures and resources. Therefore, we compared the profiles of neurologists, based within two different contexts (Profile 1 in C-O and Profile 2 in R+C). Fundamental differences were found in the quality and availability of resources, characteristics of patients and clinical needs, more focused on supporting the disease path in profile 1, whilst more orientated to discern aspects of the disease in profile 2. We also explored the contrast between two different clinical professionals, a neurologist and a psychiatrist (Profiles 2 and 3), within a similar context (R+C). In this comparison, the availability of resources, time, goals, and needs of the team are aligned, regardless of the type of specialist involved. We found that the R+C vision of a clinical context is shaped by characteristics of the clinical centre's director (who we will define as the *champion*), who might stimulate innovative ideas and a research-orientated mindset.

This analysis of user profiles gave a stronger perspective to the most suitable early adopters for our proposed tool, revealing that the type of specialisation (i.e. neurologist or psychiatrist) was not critical to the adoption of the EBM, provided that context and needs are aligned. We identified key factors that overcome these distinctions, like the presence of a champion who has a particular research-orientated vision and has access to particular facilities and resources.

4.2 Future technology adoption assessment

We then moved to framing the gap between the prospective technology and its clinical use, to assess the probability of failure or success. To meet this goal, we applied DRGm to our themes, particularly drawing on interview responses about individual practice and the reasoning around the visual stimuli (Fig. 1). DRGm was preferred to other adoption models because it is suitable for assessing a technology even at very early stages, whilst providing a semi-quantitative outcome around critical

areas for a successful adoption. Moreover, Gagnon's factors [21] used for the thematic analysis were in line with each dimension in Heeks' model (see Table 4 in Appendix A).

Q3: What are the perceived challenges and potential for adoption of the tool? Given that two main different contexts of adoption emerged from the previous analysis, we assessed both with DRGm, and the results are reported in Table 2. Each dimension provides an estimate of the actual gap between design and reality. The rating scale is derived by Hawari and Heeks [26], where 0 indicates the absence of gap, 5 is a moderate degree of mismatch, and 10 represents a profound divergence between design and reality. Ratings were assigned by the first author, who conducted the interviews and the analysis, and were reviewed by other authors. Single values were then combined to provide an overall score.

Information. The most relevant themes here referred to the EBM's *benefit* (defining stage and progression of the disease), *reliability*, and *validity*. The EBM would be beneficial to address the need for more detailed information, particularly on progression, expressed by both clinicians and patients.

> *"It's a question most patients ask: how is this going [to progress], how is the evolution going to be, what is my prognosis? [...] And at this moment we don't have any ways to predict the prognosis." (P2, C-O)*

On the other hand, a frequent concern touched on the validation and reliability of the EBM, it still being at an experimental stage. Another concerning factor was the reliability of the biomarkers' sequence generated by the model. What differentiated the two contexts is the availability of some biomarkers: in C-O clinics it is harder to access sophisticated data such as lumbar-puncture and comprehensive neuroimaging scans, whilst it is routine in R+C centres. Moreover, in R+C the majority of patients are referred for clinical uncertainty or difficult diagnosis, which shapes the information sought by the medical staff.

> *"It is easy to diagnose an evident AD. The real need is to understand more about this grey area, if we are keeping these patients." (P6, R+C)*

Technology. This dimension examined whether the current technical set-up of the workplace is suitable for the integration of the EBM. Given that the future DST can be easily implemented in desktop computers and that hospitals and clinics are already equipped with PCs, no additional technology would be needed. Notwithstanding this, data handling and interoperability need to be carefully considered.

Process. This analysis was based on the process of data collection, the workflow, and the openness to additional steps that the EBM as a tool might require (Gagnon factor *attitude*). The heterogeneity of many biomarkers that are site-dependent (e.g. cognitive tests) or asset-dependent (fluids) can be an obstacle to the use of the EBM and to a scalable tool.

> *"Certainly MMSE (Mini-Mental State Examination - cognitive test) is very rough test but very commonly used, then yes, the test might depend on the site. I think cognitive tests used are very site dependent, unfortunately. That needs standardisation." (P1, R+C)*

Regarding imaging scans, the EBM uses volumetric data as input, which is not available to clinicians. This will require an additional pipeline to the tool's back-end system.

Generally, R+C contexts have facilitated access to higher quality data and a workflow that will accommodate the future tool's usage tasks, such as the systematic recording of data in a digital database. Once the diagnosis is reached, the patient is discharged. Thus, clinicians' main interest is to follow them up in the early stages and during the eventual progression to more severe ones.

Interestingly, the tool was envisioned as a useful resource to facilitate the discussion process in MDTs:

> *"In one MDT session a multitude of cases are discussed and that implies remembering numbers of all tests, eventually compared with previous scores and other information. It is hard to keep track of all information for each." (P6, R+C)*

Objectives and values. This dimension highlighted attitudes towards the prospective tool's benefits and limitations, uncertainty, and interactions with patients. All interviewees affirmed they would require support in the diagnosis, with extra information on stages and progression. However, it is commonly recognised that the EBM would be primarily suitable for clinical trials and research settings.

> *"Where this can be very important is if you are going to do a trial for example or if you are going to pick any offered treatment for people in a particular stage, that might be helpful." (P5, R+C)*

The majority of clinicians were concerned about a precise numerical outcome for a stage, preferring a degree of uncertainty. This would allow them to combine the outcome from the EBM with their expertise. Regarding disease progression, clinicians would remain vague and hypothetical. The C-O context was more patient-focused, and concerned about how to communicate these contents to patients.

> *"We talk in vague terms about what to expect in the future, [...] we don't tend to talk very much in time spans because that's so difficult." (P2, C-O); "if you put a scale up like that [...] and say right, you've just started having memory problems, it's all very mild, that means you are at stage 5 out of 12, [the patient might reply] 'well, I think you blind me, I am half-way gone!', whereas we know that, although the disease may be present in their brain for 10 years, symptomatically they are at the beginning of that journey." (P3, C-O)*

Conversely, the R+C context was more disease-focused, highlighting the value of the EBM in advancing research, or as a support tool in daily practice.

> *"In this sense a visualisation would work. What we do now is doing that in our head, we compare with our experience, but it would be interesting to see how computers think" (P6, R+C)*

Staffing and skills. Participants reported medium/high technical skills in using electronic medical software and inspecting MRI scans, given appropriate resources and training. They found the proposed visual stimuli hard to interpret from a medical perspective.

Management. While most of the management issues are influenced by external factors (such as availability and quality of resources), the internal management structure of both C-O and R+C contexts was judged adequate to adopt a DST for the EBM, given regular quality checks and system assessment performed by a technician. Clinical champions with unique characteristics might be beneficial to encourage openness to innovation within the team, even in less suitable contexts. The implementation of the EBM can also address legal matters, as highlighted by one participant:

> *"[Staging] has a number of consequences with regard to reimbursement of medication and also to the driving capacities etcetera..." (P1, R+C)*

Other resources. The most relevant factors within this dimension were *cost* and *time-related* issues. The DST should not have a high impact on additional tasks or time-on-task. However, gathering all required data and feeding it in the tool would require additional steps that have to be carefully

Table 2. Assessment of Heeks' dimensions for the two contexts. [*critical score when equal or higher than 5.]

Heeks' dimensions	Care-Oriented	Research+Care
Information	5*	4
Technology	2	2
Process	6*	3
Objectives and Values	7*	4
Staffing and Skills	4	4
Management	3	3
Other resources	2	2
Total	29	22

Legend (likelihood as cause of failure): 0-2 very unlikely; 3-4 unlikely; 5 possible; 6-8 likely; 9-10 very likely.

considered. Costs were not explicitly assessed yet, excluding those already allocated for medical examinations.

A total score was derived from the DRGm assessment, as a sum of the scores given for each dimension (see Table 2). According to the work of Hawari and Heeks [26], the total score can be predictive of failure or success. The context related to C-O received a score of 29 (*"Project might fail totally, or might well be a partial failure, unless action is taken to close the design-reality gap"* [26]), while the R+C scored 22 (*"Project might be a partial failure"* [26]). These results reflected a smaller design-reality gap in the R+C scenario, compared to the C-O one.

5 DISCUSSION

We presented an illustrative case study of opportunities for translation and adoption of a DST built on a predictive model of disease progression called the EBM. We focussed on identifying the early adopters of a future tool and the more suitable context of use (addressed through workflows and user profiles), then explored the factors contributing to the success or failure of the novel technology [27]. We found a narrower design-reality gap in the R+C context, so this is better placed to initially adopt a future EBM support tool for AD, being mindful of the barriers highlighted by clinical experts. We will now highlight the specific barriers and facilitators in the future adoption of such tool as identified in this study, and generalised requirements for similar DSTs for AD.

5.1 Barriers and facilitators in the adoption of an EBM-based DST

In this work, the design-reality gap affected three main areas: *IT factors*, *user factors*, and *contextual factors*. IT factors are the elements of innovation proposed by the EBM and the tool; user factors are represented by the characteristics of the users, their expectations and beliefs; contextual factors indicate elements of the setting, resources, and external variables. Each of these areas contributed to define barriers and facilitators regarding the adoption of the tool considered in this study (first two columns in Table 3), as well as providing requirements for adoption of DSTs for AD in general (final column in Table 3). This last one will be further discussed in the next section (5.2).

Within *IT factors*, one barrier to adoption of a DST built on the EBM was the lack of validity and reliability. Whilst EBM has proven potential, its evidence has only been tested with research datasets. To overcome this barrier, the EBM needs to be tested in a controlled healthcare environment, and replicate its results based on a more diverse set of data. As for interpretability, a critical factor was the use of terminology. In our specific case, the EBM is programmed to detect very early stage

AD, before any symptoms occur. However, "early stages" in clinical terms are only defined when patients start to show visible symptoms, which would feature as mid-way through in the EBM staging. Therefore, this discrete numbering of stages would not suit current clinical terminology, and might also disrupt the communication with patients. On the other hand, what can facilitate the uptake of this DST is a clear and simple visualisation of the EBM output, designed based on the current terminology and data used, and that could particularly support team discussion.

User factors included attitudes, meaning the perceived usefulness of the EBM, but also the benefit gained from it. Clinicians who focus more on the disease, because of their need to assess early or uncertain cases, were more open to the advantages brought by the EBM. Some study participants were already familiar with the model, and could see aspects of it that might improve clinical practice and early detection of at-risk individuals. Whilst specialists did not find the presentation of a precise stage number useful, they suggested the inclusion of a *degree of uncertainty*. The EBM does provide this information, and including it in the visualisations will supplement the clinician's expertise.

Other implications around user factors were mostly interrelated with *contextual* ones, since users' needs and practices are influenced by their context of work [61]. The EBM tool was judged useful in MDTs, to converge perspectives from multiple disciplines towards a consensus diagnosis. This is in line with the projected key role of technology in the context of MDT meetings mentioned by Kane et al. [32]. In particular, DST in this context can foster a "shared awareness of the diagnosis" by facilitating the communication and interaction of specialists, whilst optimising the time commitment. We also noted that the EBM is not currently applicable to all contexts. We can in fact see an *early-late stage continuum*, where contexts addressing early stages of AD (such as R+C) are focused on early detection, risk of conversion, and uncertain cases, whereas needs for late stage contexts do not properly match with the EBM purposes, as they focus more on accompanying patients through their daily-life, when the condition is evident. This is an important finding, as it highlights that the EBM is likely to be more beneficial for one context than another, which will influence subsequent design decisions. However, the presence of a champion can represent an exception to these C-O contexts, e.g. by bringing a more systematic and specialised data collection approach in their team.

This study identifies and clarifies a potential value of introducing an EBM-based DST to AD clinical practice. Experiments determined that a research-oriented context is the most suitable to begin with, as the EBM will address a gap in the quantitative assessment of early stages of the disease and data-driven evolution, given multiple sources of biomarkers. However, it will be necessary to overcome current barriers to adoption related to technical, organisational and social assets. The most disruptive were the need for rigorous validation, adequate fit within context, and resource accessibility. Clarifying all these can contribute not only to guiding the research on the tool's development, but also to a higher probability of translating the EBM into a tangible technology.

5.2 Opportunities to adopt DSTs in AD clinical practice

Some of our findings can be generalised to DSTs developed for AD, but built on different underlying algorithms. This is reported in the last column of Table 3. From the IT perspective, clinicians needed transparency on DST output to feel confident in adopting them. This touches on key topics about explainability and trustworthiness of AI systems in healthcare [28, 56]. Healthcare is a particularly challenging space, where the demand for explanation is much higher compared to other applications [2], and where explanations depend more profoundly on domain knowledge and the needs of the specific context. Within healthcare, AD brings additional challenges given by its heterogeneity, which affects the range of explanations that models have to account for. Moreover, the variety of data recorded and the longitudinal characteristic of this condition makes it hard to account for

Table 3. Table summarising barriers and facilitators related to our study, and requirements for adoption of DSTs for AD.

	Barriers	**Facilitators**	**Requirements for Adoption**
IT factors	Does not use clinical terminology Lack of validity and reliability	Visualisation and team discussion	IT terminology adapted to clinical setting Data reliability and validity Interpretability
User factors	Patient-focussed clinicians Not properly fitting workflow	Disease-focussed clinicians Need to assess early or uncertain cases Familiarity with the EBM Presence of a champion	System perceived as useful System meets critical needs System should fit with workflow Team/champion promoting IT
Contextual factors	Lack of data quality C-O settings Costs	Availability of specialised data R+C settings	Availability of resources Specialised settings Costing support

consistency and data quality. Various data sharing initiatives in AD aim to generate a common platform that could guide towards data harmonisation, as a more solid base for computational models to be developed and deployed [20, 54].

Another important point refers to Human Factors, such as stakeholders' needs and beliefs [61]. Involving users early in the design process [43, 56, 57, 64] can fulfil these requirements, saving time and resources, and avoiding critical errors. Their engagement can guide in defining the intended use of DSTs. In fact, we are not only translating the outcome from an algorithm into a tool. We are actually translating a *concept* between two very different fields. To translate a concept, we need to take into account each side's beliefs, mental models, and terminology. In some cases (e.g. critical care where decisions are made under time-pressure), timing is also important. Therefore, how can we make this 'black-box' more transparent? Human Factors contribution is absolutely critical. For instance, Cai et al. [10] demonstrated that instead of improving the performance of imperfect algorithms, researchers could work on improving users' interaction with the system. This will stimulate users to build appropriate mental models through the interaction with the algorithm, and therefore gather those explanations in a progressive enquiry process.

We also stressed the importance of technology fitting within the current workflow in a seamless and interoperative way, while simultaneously bringing about positive behaviour change. This refers to the applicability or the extent to which the innovation might be suitable for a particular clinical situation (workflow, setting, needs, type of patient, and resources). One barrier to adoption was the time taken to input data in the system and who would be completing that task. Since most clinicians make use of written notes and do not record patients' data in datasets, this additional step might limit the integration of the tool into clinical routine. Other *contextual factors* included those external variables that do not depend on the user or the technology [18], such as facilitated access to data, data quality and heterogeneity, costing, and technical resources.

Our work aligns with these threads. Ours is one of the few studies [43, 61] where clinical experts were involved in the development of a clinical DST, prior to an actual tool being developed. This approach allowed us to define strategies for the future adoption of this particular predictive model in a DST for AD. Based on these outcomes, we outlined critical requirements that apply to other DSTs developed for the AD context.

5.3 Limitations and future directions

HCI research in the healthcare space is characterised by inevitable and acknowledged challenges [60], which affected this study in various ways. From a practical perspective, this included accessing clinical spaces, involving medical specialists and allocating time with them, recording protected

data, or running any tests that can interfere with doctors' routines [60]. Therefore, parts of our data collection were recorded as notes, which may be more susceptible to transcription errors than audio recordings. Moreover, our sample was not extensive, due to difficulties in recruiting suitable participants and to their limited time availability, but also to the niche target audience considered in this study. However, this number was sufficient to identify challenges, as reported above. Similar studies [9, 48, 58, 60, 61] involved small data sets to test and learn from different frameworks' applications, the sizes of which are comparable to ours. Despite this, we acknowledge that further evidence should be collected to validate our findings. From a methodological perspective, the DRGm requires the analyst to assign quantitative scoring to each of the seven dimensions. Whilst there is an inherent subjectivity in quantifying qualitative results, the real value from this conceptual contribution lies in the process that led to those scoring [46]. The scoring is presented after the description of findings for each dimension, where the reader could find critical justification for the assigned numbers. Ultimately, scoring is here intended more as a tool for the analyst to think about which gaps are more difficult to address. We have extensively used interviews for this study, which on one hand is more suitable for getting a rich input at this early stage of the development (given limited number of participants and the complex scenario), but on the other hand heavily depends on participants' ability to articulate relevant gaps in the current workflow. In particular, it is hard for prospective users to assess the real impact of a future tool on their workflow, and how they would adapt to innovation (also know as the "task artefact cycle" [11]). This is also why end-users should be kept in the loop at all stages of the development process, as they will become increasingly aware of a tool's potentials and use.

This work made progress towards a long-term goal to develop a tool that brings innovative predictive models from the EuroPOND project [17] into AD clinical practice, tailored around clinicians' needs, beliefs, and expected use. Although further engagement with a greater number of clinicians and multiple centres is needed to ensure this innovation can scale, our contribution provided the knowledge needed to confidently accomplish our next research steps, overcoming the risk of the "proximate future", so common in HCI research. By studying the barriers and potential from the real context, that could constrain the development and adoption of our future DST, we set our development goals to be totally embedded and integrated with real-world characteristics. Our future plan is to design an EBM-based DST prototype based on these requirements, and test it with the intended users to make sure we have met their needs. In the long run, and with multiple iterations, this will lead to a mature tool that can access live patient data so that a future study can test this approach "in the wild" (i.e., in ongoing clinical practice). Finally, this paper presented one example of a DST for AD; more application areas (e.g. other neurodegenerative conditions or different DPMs) should be explored, to consolidate the proposed requirements for the translation and adoption of DSTs for AD.

6 CONCLUSIONS

We explored opportunities for a novel predictive model-based DST for AD to be adopted in clinical practice, where it could advance the understanding and management of this debilitating condition. DSTs often fail when introduced to the real-world setting, due to limited human factors considerations. This is one of the few studies involving end users prior to an actual tool being developed, but it also contributed with a novel combination of methods to explore barriers and facilitators in the future adoption of such tool. There are no prior methodological guidelines on translation and adoption of a future technology in this application area, so we combined a set of methods that allowed us to tackle our key questions. We found that research-oriented centres represent a narrower gap to adoption in our case, because early adopters would be advantaged to take on the tool thanks to their facilitated access to relevant data, and their need for early diagnosis and

quantification of patients' progression that better matches the tool's capabilities. However, other internal and external factors can influence this transition, such as the validation and trust in the model, presence of a relevant champion, or characteristics and needs of the environment. Generally, the tool was perceived as having potential, if the hurdles identified can be properly addressed. This is one of the few contributions towards this topic that includes clinical users early in the process, before the implementation of the actual technology. Despite our focus on an EBM-based DST, we contributed with general requirements for other DSTs to be developed for the AD setting. Defining strategies for adoption early in the process through a user-centred approach would give timely directions for future development stages towards clinical needs. Strategies for adoption will aid future researchers in avoiding possible misconceptions, or abandonment of the technology, promoting a positive impact in real world scenarios.

ACKNOWLEDGMENTS

REFERENCES

[1] Eman AbuKhousa and Piers Campbell. 2012. Predictive data mining to support clinical decisions: An overview of heart disease prediction systems. In *2012 International Conference on Innovations in Information Technology (IIT)*. IEEE, 267–272.

[2] Muhammad Aurangzeb Ahmad, Carly Eckert, and Ankur Teredesai. 2018. Interpretable machine learning in healthcare. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 559–560.

[3] Icek Ajzen. 1991. The theory of planned behavior. *Organizational behavior and human decision processes* 50, 2 (1991), 179–211.

[4] Elske Ammenwerth, Carola Iller, and Cornelia Mahler. 2006. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. *BMC medical informatics and decision making* 6, 1 (2006), 3.

[5] Genevieve Bell and Paul Dourish. 2007. Yesterday's tomorrows: notes on ubiquitous computing's dominant vision. *Personal and ubiquitous computing* 11, 2 (2007), 133–143.

[6] Riccardo Bellazzi and Blaz Zupan. 2008. Predictive data mining in clinical medicine: current issues and guidelines. *International journal of medical informatics* 77, 2 (2008), 81–97.

[7] Amy Berlin, Marco Sorani, and Ida Sim. 2006. A taxonomic description of computer-based clinical decision support systems. *Journal of biomedical informatics* 39, 6 (2006), 656–667.

[8] Ann Blandford, Dominic Furniss, and Chris Vincent. 2014. Patient safety and interactive medical devices: realigning work as imagined and work as done. *Clinical risk* 20, 5 (2014), 107–110.

[9] Ann E Blandford, Joanne K Hyde, Thomas RG Green, and Iain Connell. 2008. Scoping analytical usability evaluation methods: a case study. *Human–Computer Interaction* 23, 3 (2008), 278–327.

[10] Carrie J Cai, Emily Reif, Narayan Hegde, Jason Hipp, Been Kim, Daniel Smilkov, Martin Wattenberg, Fernanda Viegas, Greg S Corrado, Martin C Stumpe, et al. 2019. Human-centered tools for coping with imperfect algorithms during medical decision-making. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 4.

[11] John M Carroll and Mary Beth Rosson. 1992. Getting around the task-artifact cycle: how to make claims and design by scenario. *ACM Transactions on Information Systems (TOIS)* 10, 2 (1992), 181–212.

[12] Fred D Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly* (1989), 319–340.

[13] ISO DIS. 2009. 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. *International Standardization Organization (ISO). Switzerland* (2009).

[14] Michael C Donohue, Hélène Jacqmin-Gadda, Mélanie Le Goff, Ronald G Thomas, Rema Raman, Anthony C Gamst, Laurel A Beckett, Clifford R Jack Jr, Michael W Weiner, Jean-François Dartigues, et al. 2014. Estimating long-term multivariate progression from short-term data. *Alzheimer's & Dementia* 10, 5 (2014), S400–S410.

[15] Glyn Elwyn, Isabelle Scholl, Caroline Tietbohl, Mala Mann, Adrian GK Edwards, Catharine Clay, France Légaré, Trudy van der Weijden, Carmen L Lewis, Richard M Wexler, et al. 2013. "Many miles to go…": a systematic review of the implementation of patient decision support interventions into routine clinical practice. *BMC medical informatics and decision making* 13, 2 (2013), 1–10.

[16] Arman Eshaghi, Ferran Prados, Wallace J Brownlee, Daniel R Altmann, Carmen Tur, M Jorge Cardoso, Floriana De Angelis, Steven H van de Pavert, Niamh Cawley, Nicola De Stefano, et al. 2018. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Annals of neurology* 83, 2 (2018), 210–222.

[17] EuroPOND. [n.d.]. EuroPOND Project. http://europond.eu. Accessed August 15, 2020.

[18] Louise Fitzgerald, Ewan Ferlie, Martin Wood, and Chris Hawkins. 2002. Interlocking interactions, the diffusion of innovations in health care. *Human relations* 55, 12 (2002), 1429–1449.

[19] Hubert M Fonteijn, Marc Modat, Matthew J Clarkson, Josephine Barnes, Manja Lehmann, Nicola Z Hobbs, Rachael I Scahill, Sarah J Tabrizi, Sebastien Ourselin, Nick C Fox, et al. 2012. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *NeuroImage* 60, 3 (2012), 1880–1889.

[20] Giovanni B Frisoni. 2010. Alzheimer's disease neuroimaging initiative in Europe. *Alzheimer's & Dementia* 6, 3 (2010), 280–285.

[21] Marie-Pierre Gagnon, Marie Desmartis, Michel Labrecque, Josip Car, Claudia Pagliari, Pierre Pluye, Pierre Frémont, Johanne Gagnon, Nadine Tremblay, and France Légaré. 2012. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. *Journal of medical systems* 36, 1 (2012), 241–277.

[22] Hugo Geerts, Penny A Dacks, Viswanath Devanarayan, Magali Haas, Zaven S Khachaturian, Mark Forrest Gordon, Stuart Maudsley, Klaus Romero, Diane Stephenson, Brain Health Modeling Initiative, et al. 2016. Big data to smart data in Alzheimer's disease: The brain health modeling initiative to foster actionable knowledge. *Alzheimer's & Dementia* 12, 9 (2016), 1014–1021.

[23] Conceição Granja, Wouter Janssen, and Monika Alise Johansen. 2018. Factors determining the success and failure of ehealth interventions: systematic review of the literature. *Journal of medical Internet research* 20, 5 (2018).

[24] Trisha Greenhalgh, Glenn Robert, Fraser Macfarlane, Paul Bate, and Olivia Kyriakidou. 2004. Diffusion of innovations in service organizations: systematic review and recommendations. *The Milbank Quarterly* 82, 4 (2004), 581–629.

[25] Ricardo Guerrero, Alexander Schmidt-Richberg, Christian Ledig, Tong Tong, Robin Wolz, Daniel Rueckert, Alzheimer's Disease Neuroimaging Initiative (ADNI, et al. 2016. Instantiated mixed effects modeling of Alzheimer's disease markers. *NeuroImage* 142 (2016), 113–125.

[26] Ala'a Hawari and Richard Heeks. 2010. Explaining ERP failure in a developing country: a Jordanian case study. *Journal of Enterprise Information Management* 23, 2 (2010), 135–160.

[27] Richard Heeks. 2006. Health information systems: Failure, success and improvisation. *International journal of medical informatics* 75, 2 (2006), 125–137.

[28] Andreas Holzinger, Chris Biemann, Constantinos S Pattichis, and Douglas B Kell. 2017. What do we need to build explainable AI systems for the medical domain? *arXiv preprint arXiv:1712.09923* (2017).

[29] Jan Horsky, Gordon D Schiff, Douglas Johnston, Lauren Mercincavage, Douglas Bell, and Blackford Middleton. 2012. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. *Journal of biomedical informatics* 45, 6 (2012), 1202–1216.

[30] Clifford R Jack Jr, David S Knopman, William J Jagust, Leslie M Shaw, Paul S Aisen, Michael W Weiner, Ronald C Petersen, and John Q Trojanowski. 2010. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology* 9, 1 (2010), 119–128.

[31] Monique WM Jaspers, Marian Smeulers, Hester Vermeulen, and Linda W Peute. 2011. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *Journal of the American Medical Informatics Association* 18, 3 (2011), 327–334.

[32] Bridget Kane and Saturnino Luz. 2009. Achieving diagnosis by consensus. *Computer Supported Cooperative Work (CSCW)* 18, 4 (2009), 357–392.

[33] Kensaku Kawamoto, Caitlin A Houlihan, E Andrew Balas, and David F Lobach. 2005. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *Bmj* 330, 7494 (2005), 765.

[34] Ellen Kilsdonk, LW Peute, and Monique WM Jaspers. 2017. Factors influencing implementation success of guideline-based clinical decision support systems: a systematic review and gaps analysis. *International journal of medical informatics* 98 (2017), 56–64.

[35] Hian Chye Koh, Gerald Tan, et al. 2011. Data mining applications in healthcare. *Journal of healthcare information management* 19, 2 (2011), 65.

[36] A Lanceley, J Savage, U Menon, and I Jacobs. 2008. Influences on multidisciplinary team decision-making. *International Journal of Gynecological Cancer* 18, 2 (2008), 215–222.

[37] Rosa Lau, Fiona Stevenson, Bie Nio Ong, Krysia Dziedzic, Shaun Treweek, Sandra Eldridge, Hazel Everitt, Anne Kennedy, Nadeem Qureshi, Anne Rogers, et al. 2015. Achieving change in primary care - causes of the evidence to practice gap: systematic reviews of reviews. *Implementation Science* 11, 1 (2015), 40.

[38] Choong Ho Lee and Hyung-Jin Yoon. 2017. Medical big data: promise and challenges. *Kidney research and clinical practice* 36, 1 (2017), 3.

[39] Cynthia LeRouge, Jiao Ma, Sweta Sneha, and Kristin Tolle. 2013. User profiles and personas in the design and development of consumer health technologies. *International journal of medical informatics* 82, 11 (2013), e251–e268.

[40] Christopher Lettl, Cornelius Herstatt, and Hans Georg Gemuenden. 2006. Users' contributions to radical innovation: evidence from four cases in the field of medical equipment technology. *R&D Management* 36, 3 (2006), 251–272.

[41] Joseph Lindley, Paul Coulton, and Miriam Sturdee. 2017. Implications for adoption. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 265–277.

[42] Marco Lorenzi, Maurizio Filippone, Giovanni B Frisoni, Daniel C Alexander, Sébastien Ourselin, Alzheimer's Disease Neuroimaging Initiative, et al. 2017. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer's disease. *NeuroImage* (2017).

[43] Jennifer L Martin, Daniel J Clark, Stephen P Morgan, John A Crowe, and Elizabeth Murphy. 2012. A user-centred approach to requirements elicitation in medical device development: A case study from an industry perspective. *Applied ergonomics* 43, 1 (2012), 184–190.

[44] Prabha Susy Mathew and Anitha S Pillai. 2015. Big Data solutions in Healthcare: Problems and perspectives. In *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. IEEE, 1–6.

[45] Andrew Maunder, Gary Marsden, Dominic Gruijters, and Edwin Blake. 2007. Designing interactive systems for the developing world-reflections on user-centred design. In *Information and Communication Technologies and Development, 2007. ICTD 2007. International Conference on*. IEEE, 1–8.

[46] Joseph A Maxwell. 2010. Using numbers in qualitative research. *Qualitative inquiry* 16, 6 (2010), 475–482.

[47] Kristen Miller, Danielle Mosby, Muge Capan, Rebecca Kowalski, Raj Ratwani, Yaman Noaiseh, Rachel Kraft, Sanford Schwartz, William S Weintraub, and Ryan Arnold. 2018. Interface, information, interaction: a narrative review of design and functional requirements for clinical decision support. *Journal of the American Medical Informatics Association* 25, 5 (2018), 585–592.

[48] Y Yi Mun, Joyce D Jackson, Jae S Park, and Janice C Probst. 2006. Understanding information technology acceptance by individual professionals: Toward an integrative view. *Information & Management* 43, 3 (2006), 350–363.

[49] Mark A Musen, Blackford Middleton, and Robert A Greenes. 2014. Clinical decision-support systems. In *Biomedical informatics*. Springer, 643–674.

[50] Neil P Oxtoby and Daniel C Alexander. 2017. Imaging plus X: multimodal models of neurodegenerative disease. *Current opinion in neurology* 30, 4 (2017), 371.

[51] M Prince, A Wimo, M Guerchet, GC Ali, YT Wu, and M Prina. 2015. World Alzheimer Report 2015. The global impact of dementia. An analysis of prevalence, incidence, cost & trends; Alzheimer's Disease International: London.

[52] Everett M Rogers. 2010. *Diffusion of innovations*. Simon and Schuster, 5–6.

[53] Reed T Sutton, David Pincock, Daniel C Baumgart, Daniel C Sadowski, Richard N Fedorak, and Karen I Kroeker. 2020. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digital Medicine* 3, 1 (2020), 1–10.

[54] Arthur W Toga, Scott C Neu, Priya Bhatt, Karen L Crawford, and Naveen Ashish. 2016. The global Alzheimer's association interactive network. *Alzheimer's & Dementia* 12, 1 (2016), 49–54.

[55] Samson W Tu, James R Campbell, Julie Glasgow, Mark A Nyman, Robert McClure, James McClay, Craig Parker, Karen M Hrabak, David Berg, Tony Weida, et al. 2007. The SAGE Guideline Model: achievements and overview. *Journal of the American Medical Informatics Association* 14, 5 (2007), 589–598.

[56] Niels van Berkel, Omer F Ahmad, Danail Stoyanov, Laurence Lovat, and Ann Blandford. 2020 (in press). Designing Visual Markers for Continuous Artificial Intelligence Support: A Colonoscopy Case Study. *Acm Transactions on Computing for Healthcare* (2020 (in press)).

[57] Ingrid S van Maurik, Leonie NC Visser, Ruth E Pel-Littel, Marieke M van Buchem, Marissa D Zwan, Marleen Kunneman, Wiesje Pelkmans, Femke H Bouwman, Mirella Minkman, Niki Schoonenboom, et al. 2019. Development and usability of ADappt: Web-based tool to support clinicians, patients, and caregivers in the diagnosis of mild cognitive impairment and alzheimer disease. *JMIR formative research* 3, 3 (2019), e13417.

[58] LN Visser, RE Pel-Littel, MD Zwan, M Kunneman, W Pelkmans, FH Bouwman, M Minkman, N Schoonenboom, P Scheltens, EM Smets, et al. 2019. Development and Usability of ADappt: Web-Based Tool to Support Clinicians, Patients, and Caregivers in the Diagnosis of Mild Cognitive Impairment and Alzheimer Disease. *JMIR formative research* 3, 3 (2019), e13417–e13417.

[59] Peter A Wijeratne, Alexandra L Young, Neil P Oxtoby, Razvan V Marinescu, Nicholas C Firth, Eileanoir B Johnson, Amrita Mohan, Cristina Sampaio, Rachael I Scahill, Sarah J Tabrizi, et al. 2018. An image-based model of brain volume

biomarker changes in Huntington's disease. *Annals of clinical and translational neurology* 5, 5 (2018), 570–582.

[60] Qian Yang, Aaron Steinfeld, and John Zimmerman. 2019. Unremarkable AI: Fitting Intelligent Decision Support into Critical, Clinical Decision-Making Processes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 238.

[61] Qian Yang, John Zimmerman, Aaron Steinfeld, Lisa Carey, and James F Antaki. 2016. Investigating the heart pump implant decision process: opportunities for decision support tools to help. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 4477–4488.

[62] Alexandra L Young, Neil P Oxtoby, Pankaj Daga, David M Cash, Nick C Fox, Sebastien Ourselin, Jonathan M Schott, and Daniel C Alexander. 2014. A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain* 137, 9 (2014), 2564–2577.

[63] Kai Zheng, Hilary M Haftel, Ronald B Hirschl, Michael O'reilly, and David A Hanauer. 2010. Quantifying the impact of health IT implementations on clinical workflow: a new methodological perspective. *Journal of the American Medical Informatics Association* 17, 4 (2010), 454–461.

[64] Haiyi Zhu, Bowen Yu, Aaron Halfaker, and Loren Terveen. 2018. Value-sensitive algorithm design: method, case study, and lessons. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 194.

A RESEARCH METHODS

Table 4. Gagnon's factors organised by Heeks' dimensions.

Heeks' dimensions	Gagnon's factors
Information	**1. IT related factors** 1.2 Innovation characteristics 1.3 System reliability 1.6 Evidence regarding benefits 1.7 Validity of resources
Technology	**1. IT related factors** 1.1 Design and technical concerns 1.3 System reliability
Process	**2. Individual factors** 2.2 Attitude **4. Organisational environment** 4.1 Internal environment
Objectives and values	**2. Individual factors** 2.2 Attitude **3. Human environment** **4. Organisational environment** 4.1 Internal environment
Staffing and skills	**2. Individual factors** 2.1 Knowledge **4. Organisational environment** 4.1 Internal environment
Management	**1. IT related factors** 1.5 Legal issues 1.9 Environmental issues **4. Organisational environment**
Other resources	**1. IT related factors** 1.8 Cost issues **4. Organisational environment**