Embedded Assessment of Aging Adults

A Concept Validation with Stakeholders

Matthew L. Lee & Anind K. Dey
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, USA
matthew.lee@cs.cmu.edu, anind@cs.cmu.edu

***Abstract*—Sensing systems embedded in the homes of elders have the potential to monitor individuals for early signs of functional and cognitive decline. However, it is not clear how the data collected from these embedded assessment systems can be useful for elders to support awareness and for their doctors to make better diagnoses. We conduct a concept validation of embedded assessment sensing concepts with elders, family caregivers, and clinicians. We describe their reactions to the sensing concepts, their different information needs, how they would use the information, and what limits its usefulness and provide recommendations for designers of embedded assessment systems.**

Keywords-Embedded assessment; elders; concept validation; information needs; functional abilities; sensors.

I. Introduction

Many elders experience cognitive decline as they get older. They may forget particular steps in a multi-step task or may have difficulty concentrating on a task. Occasional lapses in memory, attention, or decision-making are a normal part of aging, but consistent cognitive problems may be the first signs of progressive neurological conditions such as Alzheimer's disease or its precursor, Mild Cognitive Impairment.

Cognitive decline usually manifests itself first as changes in an individual's high-level *functional* ability, that is, the ability to carry out everyday activities. In particular, Instrumental Activities of Daily Living (IADLs) such as preparing a meal, taking medication, using the telephone, and doing housework are important for maintaining independence and require a relatively high level of cognitive ability to be performed. However, many elders [1] and even their family caregivers [2] often are not aware of the subtle changes in their functional abilities that may be early signs of progressive cognitive decline. Assessments of how individuals perform IADLs can provide early indicators for decline and allow for earlier adaptations (*e.g.*, being more careful) or interventions (*e.g.*, getting someone else to do the task) to prevent accidents, and delay institutionalization [3].

In recent years, many consumer home sensing systems (*e.g.*, [4][5][6]) have been developed to monitor the well-being of elders in their homes. Marketed with a vision of safety and security, they promise to help the elders maintain their independence and live in their own homes for as long as possible. However, these systems focus more on safety and detect high-level problems such as falls only after they occur and alert caregivers appropriately.

PERVASIVEHEALTH 2010, March 22-25, Munchen, Germany

DOI 10.4108/ICST.PERVASIVEHEALTH2010.8849

We see an opportunity for embedded sensing systems in the home to go beyond simply tracking *how often* individuals engage in IADLs by including details about *how well* they perform IADLs as earlier indicators of decline. Individuals experiencing decline usually take longer to perform these complex tasks or commit more mistakes while performing the task, even though the outcome (success or failure) of the performed task may be the same.

In this paper, we discuss the results of a concept validation study of potential embedded sensing systems designed to monitor how well elders perform everyday activities. In this qualitative study with stakeholders (elders, family caregivers, and medical clinicians), we proposed concepts for home sensing systems and investigated how these systems and the data they collect can be used to improve understanding about, and recognition of changes associated with functional and cognitive decline. We identified the information needs of stakeholders as well as what value they would gain from embedded assessment data about IADLs, including improving elders' awareness of their abilities and empowering caregivers, doctors, and occupational therapists (OTs) to make better-informed decisions for treatment. We also discuss a number of issues that need to be addressed to obtain the most value from an embedded assessment approach and provide recommendations for designers of embedded assessment systems. Before we present our concept validation, we discuss research related to our work on embedded assessment of cognitive and functional decline.

II. Related Work

Objective, timely, and ecologically valid information about the functional abilities of an individual is important for proper diagnosis and treatment of the causes of functional decline [7]. Self-report and caregiver-report questionnaires are a starting point for collecting this information, but they can be rife with biases and inaccuracies [1][2]. In the clinical setting, doctors and occupational therapists can use performance-based testing instruments (such as [8][9][10]) by having patients perform tasks in the presence of a trained observer either in the clinic or at home. The observer's goal is to detect in which low-level steps of the IADL task the patient is struggling and provide appropriate interventions. However, these assessments are expensive to conduct, as they require a trained clinician (usually an occupational therapist) to administer them in the clinic or travel to the patient's home for direct observations. Consequently, these assessments are performed infrequently and usually only *after* a problem arises. The accuracy of the results can also be biased by performance effects, where patients may act differently during the one-time assessment from how they normally function in their everyday lives. Thus, doctors need more frequent, less expensive, and more objective measures of an individual's functional ability.

The concept of *embedded assessment* of long-term functional decline of elders using automatic sensing technology in the home was first introduced by Morris *et al.* [7]. Dishman [11] also envisioned sensing systems that continuously collect data on functional abilities to promote healthy behaviors, detect diseases earlier, and facilitate informal caregiving. Our focus is on monitoring (rather than compensation or prevention) for the sake of elders, family caregivers, and doctors to gain a better awareness of changing functional abilities.

Many sensing systems have been developed that focused on monitoring *how often* IADLs such as eating, sleeping, and cooking are performed, *e.g.*, [12][13][14], with some correlating behaviors with clinical outcomes [15]. Some systems focus on recording the outcomes of a particular IADL such as medication taking [16], while others go beyond monitoring and also provide appropriate assistance, *e.g.,* [17]. They often provide binary information about whether an activity was initiated or completed.

An earlier measure of functional decline than frequency is *how well* the task was performed. The concept of *preclinical disability* [18] explains that individuals first enter a stage in which their abilities start to decline but can still maintain their functioning at a level high enough to complete the task with compensatory strategies. In this stage, neither the individual nor members of their care network realize that an underlying impairment exists nor that the individual may be on a trajectory towards disability. For example, an individual who may be experiencing decline in executive functioning abilities may have more difficulty preparing a meal that involves multiple concurrent steps. To compensate, he can slow down and pay more deliberate attention to his actions. From a purely task completion view, the outcome is the same (*e.g.*, the meal is prepared) but the underlying impairment impacts how much effort the person used to perform the task and how many recoverable errors were made during the task. Research shows that functional declines are strongly correlated with time spent on the task [10]. Thus, information about *how well* an individual performs a task – the kind of information we can obtain from embedded assessment – is important for earlier prediction [3] and treatment of decline and delaying the onset of disability as long as possible.

There have been a few attempts to design sensing systems that can not only sense *how often* but also *how well* IADLs are performed. Mihailidis *et al.* [17] developed a computer vision-based system to monitor the steps in a hand-washing task, what errors the user committed, and could also provide appropriate prompts to assist the user. Cook & Schmitter-Edgecombe [19] developed an intelligent system that can detect step errors, time lags, and missteps in the IADL task process, which can give a measure of how well the task was performed. Our work builds on this previous research in embedded assessment systems, by evaluating whether this performance information (that is, data about *how well* a task is carried out) actually provides insight to elders, family caregivers, and clinicians.

Researchers conducting participatory design of embedded assessment systems have worked with various stakeholders. Morris & Lundell [20] conducted focus groups with elders, caregivers, and doctors and identified an opportunity for technology to make more relevant and frequent assessments of functional abilities, which is the type of assessment provided by the sensing concepts we use as probes in our concept validation study. Steele *et al.* [21] found that elders were most concerned with cost and control of embedded assessment technology. Blythe *et al.* [22] found that elders valued not only safety but also their privacy and freedom, calling for "socially-dependable" design. Demiris *et al.* [23] installed sensors in the homes of elders and tracked how frequently they performed various IADLs. In participatory feedback sessions, elders said that the technology was unobtrusive and provided them with peace of mind. Rantz *et al.* [24] conducted focus groups with nurses who found interactive visualizations of how often residents engaged in various IADLs to be useful. Beaudin *et al.* [25] investigated which domains people wanted to track for long-term self-monitoring and short- term health goal tracking. These prior evaluations focused mainly on embedded assessment data about *how often* individuals engaged in different IADLs. There still remains the question about *whether information about how well IADLs were performed actually provides value to elders, caregivers, and clinicians as earlier indicators for changes in functional abilitie*s. If so, we want to understand how to present performance information to each stakeholder. Embedded assessment technologies, like many sensor-based systems, can collect an overwhelming amount of data. This raises the following questions: How can the data in low-level sensor streams be presented as salient summaries for use by stakeholders? How do the information needs of elders differ from those of their caregivers and clinicians? We will now describe our concept validation approach used to gain this understanding.

III. METHOD

We investigated these questions using a concept validation technique using concrete descriptions of our embedded assessment concepts (described below) and various representations of the data that these systems could produce. Our concept validation study was conducted with sixteen participants: four fully-functioning elders (age range 67-86), six family caregivers, three geriatricians, and three occupational therapists (OTs). We focused on independent, fully-functioning elders because they would likely benefit most from early detection. We recruited them from a social club for retired employees of a corporation. Because the caregivers of these elders lived out of town, we recruited (from Craigslist) other caregivers who looked after the health of a parent. The geriatricians and OTs worked at a large local university hospital. Concept validation sessions lasted about two hours with elders and one hour with caregivers and clinicians. Participants received $20 for their participation. First, we will describe our concept validation method, and then describe the concepts we used.

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849

The concept validation session with elders started out by asking them to assess their own functional abilities and to identify any declines in health. We discussed with them how they become aware of declines and what they do when they become aware. We then introduced three embedded assessment concepts (described in the next section) as probes for discussion to get their impressions about whether they wanted these technologies in their home. Then we showed them representations of data hypothetically generated from having these sensing concepts in their home, to probe their impressions about the usefulness of these data. We began with representations that showed the least amount of information (*e.g.*, short-term, task completion only, no process detail) and asked whether this information was useful, in what way was it useful, what action (if any) they would take, and what additional information they wanted. In response to their request for more information, we would show them other representations that had more features (longer-term views, process steps, *etc.*). Sessions with the clinicians (geriatricians and OTs) followed the same procedure but began with a discussion about how they currently collect functional data about a patient and also included a discussion about how embedded assessment systems can fit into their practice. Likewise, sessions with caregivers began with asking them how they currently keep track of their parents' health. Based on transcribed audio recordings of the concept validation sessions, we used grounded theory [26] to code each transcribed comment from our participants and generate themes common across our stakeholders. Stakeholder comments about each data dimension were identified and grouped. Now we describe the sensing concepts we used as probes in our concept validation study.

A. *Sensing Concepts*

We generated the following sensing concepts for embedded assessment of specific IADLs: Medication Monitor, Coffee Chronicler, and Telephone Tracker. Medicine taking, coffee making and telephone use were chosen based on a number of factors. We considered the entire canonical list of Instrumental Activities of Daily Living because they are commonly used in existing self-report, informant report, and expert assessment instruments [8][9][10]. We also considered the current state of sensing technology so that our concepts would be feasible for implementation. We also observed how elders perform these tasks in their everyday routines to identify the low-level steps and to understand how existing simple sensors could detect the individual steps of the tasks.

Due to the fact that different individuals carry out the same IADL differently, the particular sensing system implementation must be customized for the particular way an individual carries out the task. In fact, [7] called for "extreme personalization" of sensing to reduce intrusiveness and to allow an individual to leverage her unique system of artifacts, reminders, and tools currently used to carry out the task. The actual physical sensors may vary across individuals, but the aggregated data they collect will correspond to the same higher-level task milestones (*e.g.*, selecting the right pill can be sensed with either a camera, a smart pillbox, or RFID-tagged bottles.) Well-designed standards for sensor data can help reduce the costs and complexity of interpreting diverse sensor data. We customized our sensing concepts based on the approach used by one particular individual from our observations and had elders imagine that they carried out the steps according to this representative individual. (We found that our scenarios' task steps were very similar to those of our participants who were able to adequately relate to the scenarios.) We excluded the implementation details of the proposed devices because we wanted to evaluate the value of the concepts rather than the hardware. We chose the following three sensing concepts to use as probes.

1) *Medication Monitor*

The Medication Monitor concept was customized for one particular method of taking medications used by one of the elders we observed in our preliminary observations. This individual retrieves all her daily pills from the appropriate slot of her pill box located in her kitchen. She then places all the pills onto the kitchen table and sorts out which to take now and leaves the remaining pills on the table. She then retrieves her dedicated water glass on the counter, fills it with water from the tap, and takes her pill with water. We consulted OTs to ensure that this process is common among elders. The Medication Monitor consists of a smart pillbox, a vision-enabled kitchen table, and an augmented water glass. The smart pillbox knows its location, when the user is grasping it, which doors are opened, and how much time the individual takes to decide which door to open. Once the pills are placed on the table, the vision-enabled kitchen table uses a ceiling-mounted camera to identify which pills are on the table and to monitor the pill-sorting task. The intelligent water glass senses its position on the counter, when it is filled, when it is grasped, and when it is tilted while drinking. The combination of these various devices can be used to sense when each step is started or finished, how long she spends in each step, and when errors occur. For example, taking a longer time to sort through the pills on the table could be an indicator for confusion that may eventually lead to taking the wrong pill. Leaving the pills on the table instead of taking them might show the individual is forgetting to complete all the steps. The Medication Monitor concept is similar to previous systems such as [16] for tracking medication-taking but goes beyond the pill box/bottle and looks at the entire process of taking medication.

2) *Coffee Chronicler*

The Coffee Chronicler detects the sequence of steps required for making a pot of coffee, which were fairly consistent across our participants. The concept consists of an augmented coffee maker that can detect if the carafe is empty, the quantity of coffee grounds in the machine, how much water is the machine, and whether the ratio of coffee to water is reasonable. Not only can this concept detect how often the user engages in making coffee, but it can also detect when steps are missed, repeated, or performed not as well as they should be (for example, measuring out too many scoops of coffee grounds). Other errors can include forgetting to put in the water, putting coffee in the machine before replacing the filter, or not turning on the machine.

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849

3) Telephone Tracker

The Telephone Tracker monitors the frequency of incoming and outgoing calls, which may provide an indicator of social connectedness. The Telephone Tracker is not only able to track if calls were made successfully but also can detect the errors in the task process. For example, it can detect when the user dials a wrong number by recognizing a pattern when the user dials a number, listens for a short while, hangs up, and dials again, possibly a number very similar to the first number dialed. The concept can also track how long it takes for the user to recognize the caller on the other end of the line by measuring the time between the caller introducing himself and the user responding with an appropriate greeting.

B. *Data Representations*

Based on our sensing concepts, we generated simulated data that would be collected if these systems were deployed for a year in an individual's home. Our data representations of IADL task behavior showed three features made possible by the low-level sensing of the tasks [Fig. 1, 2, & 3]: 1) task performance (instead of only task completion), 2) long-term view, and 3) process details about individual steps of the task.

1) Task Performance

Like other related systems, our embedded assessment concepts can sense whether an individual has completed the task or not. However, our concepts were also designed to track *how well the user performs these tasks*. Included in the measure of task performance are: the amount of time spent on the task, how accurately they performed the task (*e.g.*, measuring out coffee), and the number of recovered errors. One of our simulated data examples (Fig. 1) showed nearly perfect task completion early on (*e.g.*, no missing pills) but, at the same time, also showed inefficiencies involved in the task (*e.g.*, taking longer than usual to sort the pills). We explored whether stakeholders could understand the idea of task performance and find it useful beyond knowledge of task completion.

2) Long-term and Short-term View

Our representations either showed a longitudinal range of data (a year's worth of aggregated or sampled short-term data, *e.g.*, Fig. 1) or short-term data (*e.g.*, Fig. 2) that shows task status for a single day or week. Many home sensing systems emphasize intervention based on short-term data about task completion, so we wanted to assess the value in viewing long-term data about task performance.

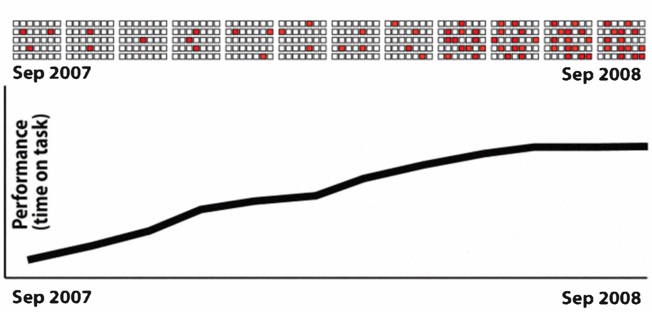

Figure 2. Long-term view of data, showing both task completion (shaded square means non-completion with medication task on that calendar day) and task performance (shown as time-on-task in line graph). The task performance data show a preclinical stage of disability.

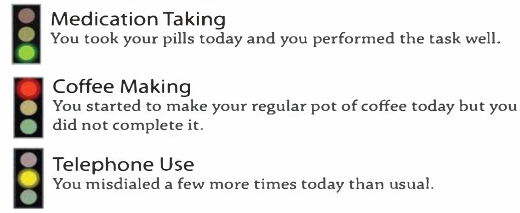

Figure 1. High-level data representation that shows short-term task performance. A green light indicates normal performance, a yellow light indicates decrease in task performance for the current day, and a red light shows a failure in task completion.

3) Process Details

One of the fortunate side effects of designing a system that monitors the task performance in addition to merely task completion is that the system has intimate knowledge of each atomic step in the process of carrying out the IADL. We provided information (*e.g.*, Fig. 3) about which steps were completed, attempted but not completed, not initiated at all, or completed out of order. We investigated whether this highly-detailed information would be useful for understanding the precise nature of any breakdowns observed and for developing appropriate interventions.

IV. Results

In this section, we discuss the results of our concept validation by first describing how embedded assessment data provides stakeholders with a greater awareness of changes in functional abilities and then describing what specific features of the data were valuable to different stakeholders and how these features would support their goals.

A. *Increased Awareness of Functional Abilities*

During our interviews, the elders without any significant functional deficits in our study experienced a conflict between their current sense of awareness of their own abilities and their concerns about losing awareness in the future. In their current state, a monitoring system is redundant because they feel they know and can stay aware of their own capabilities, breakdowns, and inefficiencies. Many said they saw the need for monitoring only after they start to have a problem with these particular tasks. For example, Elder #1 (E1) said, "*If it got to the point where it was essential … where I was making mistakes, then some sort of a system like that would be useful*."

However, the same elders also recognized that they may lose their ability to stay aware of changes in their abilities.

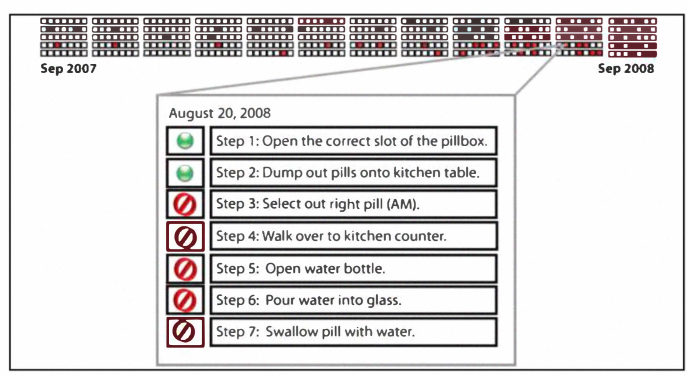

Figure 3. Process detail for an individual's medication-taking task. It shows that the individual completed the first two steps but did not attempt the last five steps.

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849

Many reflected on the experiences of their own parents or older friends as they struggled with decline in the last stages of their lives. E3 said, "*We know when we get older we will lose our hearing and memory.*" E2 said "*I want the information in my house when data indicates that people around my age or a little younger start to get problems that are serious...I don't want to wait until I know there is a problem.*" As a result, these elders expressed a desire to have embedded assessment in their homes right now so they can maintain awareness and remain functional longer. Even though we did not include the perspectives of more impaired individuals in this study, these perspectives came through the vicarious experience of the healthy individuals.

Elders had no problems and even suggested sharing detailed information about private IADLs, such as medication taking, with their family, doctor, and close friends. E3 said, "*We talk about our health problems all the time with our friends anyway.*" Clinicians were more concerned about privacy and made it clear that patients must know that they are being monitored before the data could be used in the clinic.

The geriatricians in our study indicated that embedded assessment data provides them with information they do not normally have access to, especially due to the limited amount of time (a few minutes) they can spend asking patients about the details of their abilities. OTs are accustomed to dealing with functional assessment data but said embedded assessment data could provide them with a larger time window into a patient's abilities rather than infrequent snapshots of functioning. Caregivers also found the embedded assessment process data to be useful for showing details about their loved one's abilities that they would not normally know because they do not normally talk about these types of "mundane" tasks. Caregiver #6 (CG6) said, "*I don't know if she's screwing up and not telling me. She could be screwing up everyday making her coffee. It would show that she was slipping more than she would let on.*" In the following section, we describe which specific features of the data each stakeholder wanted, which can help inform the design of representations of embedded assessment data.

B. Information Needs of Elders, Caregivers, and Clinicians

We identified the information needs of embedded assessment stakeholders which included elders, their family caregivers, geriatricians, and occupational therapists. Here, we describe which features were desirable for each stakeholder.

1) Task Completion vs. Task Performance

All stakeholders found the task completion information (whether the task was completed with an acceptable outcome) useful because it showed how often the individual did not or was unable to complete the task. When an important task such as medication-taking was missed consistently, all stakeholders recognized the need for intervention. In addition and more importantly, all stakeholders also found task performance (the quality of the outcome, the amount of effort or time spent, the number of errors encountered during the task) to be helpful.

OTs found performance time and the number of errors to be valuable in their practice because it gives them a measure of *adequacy*, a measure they normally look for in functional assessments. To apply the appropriate adaptation, OTs evaluate functional abilities along three different criteria: *independence* (the ability for the patient to carry out the task on their own), *adequacy* (the ability of the patient to perform the task with precision and economy of effort), and *safety* (the ability of the patient to avoid potentially dangerous situations) [9].

Elders also said that task performance data would provide them with early indicators for problems. Regarding Medication Monitor data, E2 said, "*[Task completion] is more useful for telling you that you're not taking your pills, [task performance time] is more useful for telling you what the problem is.*" Elders used these early indicators as triggers for adaptations to ensure task completion. For example, when seeing an increase in misdials (while still completing the call eventually) in the Telephone Tracker data, E4 said, "*It helps me understand what the problem is and how I can correct it. Gee, I'd better take a little more time in dialing, or make the names bigger.*"

Geriatricians said that performance data provided them with more information to understand the patient's abilities from a qualitative standpoint than they could get from a clinical test or observation. Geriatrician G3 said, "*Absolutely, it's a trigger for clinical investigation, to bring the patient in, sitting down and talk and figuring out what's going on.*" Caregivers said they would use decreases in task performance as a trigger to keep a closer eye on their loved one and to start a conversation if the changes are getting worse. CG3 said, "*I think it's because [it] initiates a conversation between me and my mother. I would never know that she was misdialing more often unless she complaining about it.*"

2) Long-term vs. Short-term

All interviewees found the long-term view of the data to be useful for understanding the trajectory of decline. Geriatricians said that the long-term view provided them with information about the evolution of the disability. A sudden onset of a problem can indicate an acute (or even temporary) change due to some trauma or change in the patient's life. A gradual onset of a problem can indicate a pattern more consistent with a progressive disease such as dementia. OTs found the long-term view useful for understanding the nature of the particular disability and identifying the variation in people's abilities over time. OT2 said, "*An assessment visit is a one-shot deal. It's almost impossible to observe people's habits.*"

Elders also considered the long-term representation useful for understanding how their abilities change. In fact E4 said that the long-term view "*sells the entire idea,*" meaning that the longitudinal data showing a pattern of decline provides a compelling reason to have an embedded assessment system in his home. Unlike geriatricians and occupational therapists, elders also said they would also like short-term views of the data, particularly for giving them an extra sense of security for the memory-intensive task of taking medications, which is consistent with findings from [27]. E2 said he wanted to have a more immediate indicator for missed pills or pills taken twice than his system of writing the date on the pill bottle. For the coffee making and telephone task, elders said they would be able to remember and notice if they had problems because they have more noticeable, though less critical, consequences, so short-term information about these tasks would be unnecessary.

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849

3) Process Details

Elders, caregivers, and occupational therapists were interested in the breakdown of tasks at the process level because these stakeholders have the responsibility to identify and fix problems. In contrast, geriatricians found the process steps information to be too detailed. Geriatricians pointed to social workers, OTs, or geriatric nurses as professionals better suited for acting on process information. G2 said, "*I don't think it's appropriate for me to do all the things that other team members do*." In their regular practice, OTs decompose tasks into individual steps so that they can identify the exact disability and provide the most appropriate adaptation to compensate for that disability. Caregivers wanted initiate a dialog to find the causes of the problem and a solution to help their loved ones to remain independent. CG2 said process details were helpful "*because you can know what pills are giving her the problem. The problem is to figure out the right solution [and] how to make it easier for them.*" Most elders also felt the process details were helpful in finding the exact problem to address. However, one elder, E4, said that this information was overkill because reading all the steps and seeing which were good or bad was too cognitively demanding and perhaps emotionally alarming. He said his ability and patience to look through these reports would be even less as his abilities decline. E4 said, "*I like the data but I don't think I can make sense of it. Having a little voice come out and say 'It's taking you a lot of time to do step one' that would be very helpful.*" He suggested the system highlight only the salient deviations instead of every step.

V. Discussion

In our validation sessions, stakeholders also brought up some limitations in being able to interpret the data captured by the sensing concepts. These limitations discussed below lead us to design recommendations (Table 1) that will help inform the design of embedded assessment systems that can assist stakeholders in making sense of task performance data.

A. The "Why" is missing from the data

Embedded assessment holds the promise that technology will be able to collect real life data from users to provide the answers to many of the questions doctors have about the deficits that people encounter in their everyday lives. However, our investigation revealed that the data collected from our sensing concepts are merely observed behaviors that require further explanation. An observed behavior can have any number of bio-psycho-social causes. For example, when looking at a chart showing that a patient has been increasingly skipping his daily medication, G2 remarked that while it shows an important pattern, the chart does not show why this behavior occurred. She said she would "*think about whether the [data] is providing information in a way that makes you think about different reasons for non-adherence.*" These reasons can include cognitive problems (*e.g.*, forgetting to take their pills), medical problems (*e.g.*, avoidance due to an unpleasant side effect), psychological problems (*e.g.*, depression), or financial problems (*e.g.*, can no longer afford to purchase medication). Doctors would also like to be made aware of a common deficit that manifests across different tasks. For example, a memory deficit might show up as beginning the pill taking routine but not completing it, forgetting to turn off the coffee maker, and dialing out-of-date phone numbers. The data should be shown in a way that makes these associations easy to spot.

Geriatricians said that they would engage the patient and their relatives in an extended interview and ask about their awareness of specific trends found in the embedded assessment data, to identify the possible reasons for the trends and provide the appropriate treatment. Likewise, when presented with the data about task inefficiencies or errors, caregivers would call up their loved one to find out the causes of the behaviors and try to assist them. CG5 said, "*I'd probably talk to them about it. Try to troubleshoot and see if their routine had changed.*"

Occupational therapists, with their perspective of restoring functional abilities by intervening with compensatory techniques, need to know both the problem and its causes to apply the right adaptation. For example, consider two causes for skipping medication: forgetting to take the pills (cognitive) and not being able to reach all the pills in the pillbox (dexterity). An OT would apply two different adaptations to support the task: moving the pillbox to a more noticeable position or replacing the pillbox with a larger pillbox that is more easily grasped, respectively.

Elders expressed a need to understand the reasons for changes in their health as they get older. E3 remarked, "*As an individual, we like to figure out why.*" Embedded assessment data triggers them to investigate the causes and take proactive steps to control problems before they become bigger problems. For example, in reaction to data showing an increasing number of telephone misdials, E1 said he would like to know which numbers he was misdialing so that he could figure out what might be causing this behavior and stop misdialing.

We observed that embedded assessment data are best used as a trigger to explore and address the underlying causes of the problematic behaviors, rather than providing conclusive answers about the exact causes of the behavior. Thus designers working with this data should enable the user to investigate why the problems occurred. For example, systems can arm clinicians or caregivers with specific questions driven by the data to ask during a visit, or visualizations integrating multiple streams of data may help reveal the same underlying deficit.

B. Searching for Significance

Because the stakeholders have never been presented with the fine-grained and frequent data points provided by embedded assessment technology, they had difficulty determining when the illustrated changes in performance were significant enough to warrant concern and further action. Geriatrician G2 said about the data from the Telephone Tracker concept, "*At this level of subtlety, I probably don't know how much misdialing the patient would have to do before I would… do my cognitive impairment screening.*"

Geriatricians wanted to use these data in their clinical practice but expressed the concern that they needed a way to standardize the interpretation of the data. A normal part of their methodology is quantifying people's abilities and disabilities so that they can consistently and confidently apply heuristics (such as the DSM-IV [28] criteria for dementia) for medical diagnoses of decline. Traditionally, doctors would rely on the

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849

subjective self-reports and reports from relatives about the functional abilities of the patient to provide them with evidence about how the impairment is interfering with everyday life. Now equipped with objective embedded assessment data, doctors want to operate on this objective data in a quantitative manner similar to how they operate on objective cognitive testing data such as from the Mini Mental Status Exam [29]. Doctors wanted embedded assessment data to be validated by their community so they can apply a heuristic (such as "*If the patient misses their medication 25% of the time, then it's time to be concerned and figure out what's going on.*"- G1). Occupational therapists also called for these task-specific "critical values" in the embedded assessment data to signal when these failures are interfering with the life of the patient. OT1 said, "*There would have to be critical red flags. How many times do they forget a certain step that's critical…by the tenth time or the sixth*?" The elders expressed the same need to understand when a change in observed functioning is sufficiently severe as to warrant either a minor reaction such as extra vigilance or a major reaction such as scheduling an appointment with a doctor or considering moving into an assisted living facility.

Caregivers on the other hand were able to decide on what data values would trigger them to initiate a conversation or provide assistance. The threshold values varied across different caregivers and was mostly determined by the caregiver's relationship with the individual. Some caregivers who keep in close contact with their loved one would ask about any small change, whereas others wanted to minimize their own intrusiveness into their loved one's life and would react only when they saw a dramatic decline in abilities. Even though some caregivers were hesitant to react to small changes, they still emphasized that they wanted to know about the small changes including the task performance information to understand how their loved one is doing on a regular basis. CG6 said, "*Seeing this line [Fig. 1], it would indicate to me to keep a better eye on it.*"

Another factor that contributes to searching for significance was that some tasks are easier to determine critical values for than other tasks. For instance, all stakeholders easily set critical values for medication taking to be very low such that almost any change in performance warrants some investigation...In contrast, the coffee and telephone tasks were less critical for safety or health, so the critical values for the number of errors, missteps, or misdials are higher and less well-defined.

Embedded assessment systems should initially monitor tasks that have easily-defined critical values and should also closely align their approach with standardized functional assessments. Data from embedded assessment systems need to be correlated with other well-established outcome measures such as psychometric tests or diagnoses of dementia. Evaluations of embedded assessment systems should include measures for clinical outcomes. Evaluations such as in [15] and [30] provide good examples to follow.

C. Noisy Data from the User, Not from the Sensors

Even in the world of perfect sensors that can accurately detect people's actual behaviors, people's performance of tasks can be (and will likely be) highly variable. Unlike many applications of sensing technologies, embedded assessment not only has to deal with the noise generated from the sensors themselves but also the variability in the underlying behavior being sensed. Geriatrician G2 noted that many individuals do not follow a smooth, predictable stage of preclinical decline in functioning before the onset of a disease or dementia. People may experience a decline, recover momentarily, and revert back to a pattern of decline *or not*. The fact that embedded assessment technologies can capture performance data frequently at a high level of detail makes the temporary changes (potentially noise) in performance more apparent in the data. G2 remarked, "*There's just phenomenal variability in performance. People don't decline in a steady fashion. They're waxing and waning all over the place.*" Although the high-resolution data may be noisy, it can still enable clinicians to see a change from consistency to volatility in performance which is predictive of future decline.

Even if their abilities are relatively stable, individuals may still occasionally deviate from their routine when it is convenient to do so. For example E1 said, "*I might not take all my pills all at one time, I might get up early, take a glucosamine [pill] if my knees hurt and then sleep until 11 and then take the rest then.*" No stakeholders wanted these small deviations to be flagged as errors because they are considered as "acceptable" noise. Some recommendations for designing sensing systems to accommodate noisy behaviors include: 1) Building rich models of users' actions including edge cases, and 2) Making sensing systems easy to update to accommodate acceptable deviations from established patterns.

The promise that embedded assessment will automatically provide early detection of disability based on clear, steady trends in the data may be more difficult to achieve than previously thought due to large variability in the actual behaviors being sensed. Geriatricians said even they have problems identifying meaningful patterns from the noisy data, so it would be difficult to automate this. Clinicians felt that the system should refrain from making a medical interpretation of the collected data, but rather allow clinicians to use their own experience and insights to figure out what problem(s) exists and exactly what caused it. Clinicians were comfortable with having systems take the role of identifying clear statistical patterns within variations and even suggesting particular avenues of inquiry. Embedded assessment systems can present information, highlight relationships, and even suggest causes but they should not aim to replace the clinical judgment.

VI. Conclusion & Future Work

Embedded assessment technology can be used to monitor *how well* elders perform IADLs, not merely whether they were completed or not. Information about task effort, accuracy, and errors provide early indicators of decline before actual failures in task completion occur. Our subjects found this information from embedded assessment systems to be valuable, in providing increased awareness for elders and their family caregivers, and facilitating clinical judgment for geriatricians and occupational therapists. They found great potential in our concept sensing systems and the idea of embedded assessment.

However, we also discovered a number of issues that will impact how we will actually construct and deploy instances of

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849

Table 1. Summary of Findings and Design Recommendations

Findings	Design Recommendations
Embedded assessment increases awareness and is useful in clinical judgment.	Provide appropriate representations of data to different stakeholders to support awareness for elders and provide ecologically-valid, longitudinal data to clinicians.
The "Why" is missing from the data.	Embedded assessment data are merely triggers for further explorations of underlying health issues. Support data-driven inquiry with the user.
Lack of validation for critical values of significance.	Codify embedded assessment data into scales with critical values to quantify significance in the data. Correlate embedded assessment data with standard psychometric instruments.
The underlying behavior is noisy, not just the sensor data.	The system should include a rich model of the user's actions to accommodate acceptable deviations from established routines. The system can highlight patterns and make simple suggestions to facilitate data exploration but should ultimately allow clinicians to use their own experience and intuition.

embedded assessment in the next phase of our work. We found that different people carried out IADLs very differently from others, and thus embedded assessment systems that track task effort must be customized to the particular individual's method of carrying out the task. Our concept validation revealed three important issues that limit the usefulness of these systems. The data they produce do not explicitly explain why particular behaviors were observed, so sensing systems should either highlight the abilities underlying these behaviors or provide its data as triggers for further investigation. There is currently a lack of any standardized metrics by which to identify what frequency or severity of problems behaviors is significant enough to require action. Therefore, future evaluations of embedded assessment technology need to be correlated with functional clinical outcomes. Embedded assessment technology must not only deal with the noise in the sensing devices but also the large variations in the performance of IADLs. Systems can perform the statistics and highlight trends but should rely on the expertise of the user to make sense of the data.

Based on our findings from this work, we plan to design and implement an embedded assessment system to monitor critical IADLs and deploy them in the field to correlate them with validated measures of cognitive and functional decline. As a formative evaluation method, concept validation is limited in that it does not use real data, and thus we will perform further longitudinal validation of the value of embedded assessment for increasing awareness of functional decline and providing early signs of degenerative cognitive conditions like dementia.

VII. REFERENCES

[1] O. Okonkwo, R. Griffith, D. Vance, D. Marson, K. Ball, and V. Wadley, "Awareness of Functional Difficulties in Mild Cognitive Impairment: A Multidomain Assessment Approach" J Am Geriatr Soc, vol.57, pp.978-84, 2009.

[2] N. Kemp, H. Brodaty, D. Pond, and G. Luscombe, "Diagnosing Dementia in Primary Care: The Accuracy of Informant Reports" Alzheimer Disease & Associated Disorders, vol. 16, pp.171-176, 2002.

[3] P. Barberger-Gateau, C. Fabrigoule , C. Helmer , I. Rouch and J. Dartigues "Functional impairment in IADLs: an early clinical sign of dementia?" J Am Geriatr Soc, vol. 47, pp.456-462, 1999.

[4] Eaton Home Heartbeat. http://www.homeheartbeat.com/

[5] HealthSense eNeighbor. http://www.healthsense.com

[6] SimplyHome. http://www.simplyhome-cmi.com/

[7] M. Morris, S. Intille, and J. Beaudin. Embedded Assessment: Overcoming Barriers to Early Detection with Pervasive Computing. *Pervasive*: 333-346, 2005.

[8] M. Diehl, M. Marsiske, A. Horgas, A. Rosenberg, J. Saczynski and S. Willis, "The revised observed tasks of daily living: A performance-based assessment of everyday problem solving in older adults" Journal of Applied Gerontology, vol. 24, pp.211-211, 2005.

[9] M. Holm and J. Rogers, "Performance assessment of self-care skills. Assessment in occupational therapy mental health: an integrative approach." B. Hemphill-Pearson. Thorofare, NJ, Slack Inc, pp.117–24, 1999.

[10] C. Owsley, M. Sloane, G. McGwin Jr, and K. Ball, "TIADL: Relationship to cognitive function and everyday performance assessments in older adults" Gerontology vol. 48, pp.254-265, 2002.

[11] E. Dishman, "Inventing wellness systems for aging in place" *Computer* vol. 37, pp.34-41, 2004.

[12] M. Philipose, *et al.*, "Inferring Activities from Interactions with Objects." IEEE Pervasive Computing vol. 3, pp.50-57, 2004.

[13] M. Stikic, T. Huynh, K. Van Laerhoven, B. Schiele, "ADL recognition based on the combination of RFID and accelerometer sensing," In Proc. PervasiveHealth 2008, pp.258-263.

[14] E. Tapia, S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors" Lecture Notes in Computer Science, vol. 3001, pp.158-175, 2004.

[15] J. Kaye, *et al.*, "Deploying wide-scale in-home assessment technology." In Proc. International Conference on Technology and Aging, 2007.

[16] T. Hayes, J. Hunt, A. Adami, and J. Kaye, "An Electronic Pillbox for Continuous Monitoring of Medication Adherence." In Proc. IEEE Engineering Medicine and Biological Society, pp.6400-3, 2006.

[17] A. Mihailidis, J. Boger, M. Canido, and J. Hoey. "The use of an intelligent prompting system for people with dementia" Interactions, vol.14 pp.34-7, 2007.

[18] L. Fried, S. Herdman, K. Kuhn, G. Rubin, and K. Turano, "Preclinical Disability: Hypotheses about the Bottom of the Iceberg." J Aging Health, vol. 3, pp.285-300, 1991.

[19] D. Cook, and M. Schmitter-Edgecombe, "Assessing the quality of activities in a smart environment" Methods Inf Med, vol.48, pp.480-5.

[20] M. Morris and J. Lundell, "Ubiquitous computing for cognitive decline: Findings from Intel's proactive health research", Intel Corporation, 2003.

[21] R. Steele, C. Secombe, W. Brookes, "Using Wireless Sensor Networks for Aged Care:The Patient's Perspective," In Proc. PervasiveHealth 2006,

[22] M. Blythe, A. Monk, K. Doughty, "Socially dependable design: The challenge of ageing populations for HCI" Interacting with Computers, vol. 17, pp.672-689, 2005.

[23] G. Demiris, *et al.* "Nurse Participation in the design of user interfaces for a smart home system" Conference on Smart Homes &Health Telematics, 2006.

[24] M. Rantz, M. Skubic, S. Miller, and J. Krampe, "Using technology to enhance aging in place" Lecture Notes in Computer Science, vol. 5120, pp.169-176, 2008.

[25] J. Beaudin, S. Intille, and M. Morris, "To Track or Not to Track: User Reactions to Concepts in Longitudinal Health Monitoring" J Med Internet Res, vol. 8, pp. e29, 2006.

[26] J. Corbin and A. Strauss, "Basics of qualitative research: Techniques and procedures for developing grounded theory" SAGE Publications Ltd, 2007.

[27] Y. Lee, J. Tullio, N. Narasimhan, P. Kaushik, J. Engelsma, S. Basapur, "Investigating the potential of in-home devices for improving medication adherence," *In Proc. PervasiveHealth 2009* pp.1-8.

[28] Diagnostic and statistical manual of mental disorders (DSM-IV). (1994). Washington, D. C. A. P. A.

[29] M. Folstein, S. Folstein and P. McHugh, "Mini-mental state. A practical method for grading the cognitive state of patients for the clinician" Journal of Psychiatric Research, vol.12, pp.189–98, 1975.

[30] H. Tyrer, M. Aud, G. Alexander, and M. Skubic, "Early Detection of Health Changes In Older Adults" Engineering in Medicine and Biology, 2007.

Digital Object Identifier: 10.4108/ICST.PERVASIVEHEALTH2010.8849
http://dx.doi.org/10.4108/ICST.PERVASIVEHEALTH2010.8849