

Letting Patients' Daily Living Information Speak: A Novel Approach to Study Geriatric Patients with Dementia and Hypertension

Weifeng Xu, William R. Betz, Stephen T. Frezza, and Wookwon Lee

Abstract— Geriatric patients with dementia and hypertension (DAH) suffer both physically and financially. The needs of these patients mainly include improving the quality of daily living and reducing the cost of long-term care. Traditional treatment approaches are strained to meet these needs. The goal of the paper is to propose an innovative treatment approach to provide cost-effective quality therapies for geriatric patients with DAH by collecting and analyzing the multi-dimensional personal information, such as observations in daily living (ODL) from a non-clinical environment. The proposed ODLs in paper include activities, cleanliness, blood pressure, medication compliance and mood changes. To complete the system design, an incremental user-centered strategy is exploited to assemble needs of patients, caregivers, and clinicians. A service-oriented architecture (SOA) is employed to make full use of existing devices, software systems, and platforms. This health-related knowledge can be interpreted and utilized to help patients with DAH remain in their homes safely and improve their life quality while reducing medical expenditures.

Index Terms— Geriatric patients with dementia and hypertension, observations in daily living, decision support, health care system, service-oriented architecture.

I. INTRODUCTION

Currently, over five million Americans have dementia and the number is expected to triple by the year 2050 [1]. Furthermore, fifty-two percent of patients with dementia also have hypertension, which is the most common coexisting medical condition associated with dementia [2]. Loss of functioning complex tasks of everyday life is a hallmark feature of the geriatric patients with dementia and hypertension (DAH), such as cognitive decline as well as loss of competence in either social or occupational domains [3], [4]. The proposed project seeks to exploit current information technologies, including theories, algorithms, processes and tools, to develop a system for testing whether and how multi-dimensional personal information collected from a non-clinical environment, referred to as the observations in daily living (ODL) [5], can provide new insights into the geriatric patients with DAH.

Manuscript received July 30, 2010.

W. Xu is with the Computer and Information Science Department, Gannon University, Erie, PA 16501 USA (phone: 814-871-5708; fax: 814-871-7616, e-mail: xu001@gannon.edu).

W. R. Betz is with National Health Information Services, Erie, PA 16506 USA (e-mail: willbetz@ntlhealth.org).

S. T. Frezza is with the Computer and Information Science Department, Gannon University, Erie, PA 16501 USA (e-mail: frezza001@gannon.edu).

W. Lee is with the Electrical and Computer Engineering Department, Gannon University, Erie, PA 16501 USA (e-mail: lee023@gannon.edu).

Towards achieving this goal, the objectives of the proposed approach are 1) to design and implement an integrated system to collect and transform multi-dimensional personal data, 2) to test and analyze the usability and reliability of the data collection and transformation processes, 3) to systematically evaluate how the proposed system responds to the needs of DAH patients, caregivers, and clinicians based on the knowledge derived from the multi-dimensional personal data, and 4) to integrate the proposed system into clinical practice workflow and evaluate the anticipated impact on clinical practice.

In the proposed project, we adopt a user-centered strategy to assemble needs of patients, caregivers, and clinicians. We have identified a range of ODLs for gaining a greater understanding of DAH patients' symptoms. To complete the architecture design, a service-oriented architecture (SOA) is employed to make full use of existing devices, software systems, and platforms [6]. For instance, we have adopted the iPhone 2.0 software development kit (SDK) as a communication platform to transmit health-related information, such as ODLs, to remote servers. We also adopt the Google Health platform to store patients' personal health profiles. The published iPhone and Google Health APIs strongly support the data integration process, which combines the patient health profiles and ODLs as a whole. The combination of patient health profiles and ODLs is an essential for interpreting personal health information in a non-clinical environment. Adopting SOA increases productivity and improves quality while reducing the total development costs of the proposed system.

II. VISION AND HEALTH IMPACT

The vision of our project is to help geriatric patients with DAH, their caregivers, and clinicians to discover and understand health-related knowledge obtained from the proposed multi-dimensional ODLs in a non-clinical setting and incorporate such knowledge into their beliefs, values, procedures, and actions for improving the daily living quality of these patients while reducing the cost of long-term care.

The core of addressing the illness concerns of the geriatric patients with DAH is to utilize IT management solutions to manage personal health information. As shown in the left triangle in Fig. 1, the personal health information in our system is categorized into three levels: 1) the illness concerns of the target patients, such as living independently and safely for DAH patients, 2) the symptoms of DAH, and 3) the ODLs

associated with the DAH symptoms selected for our project. Each level of personal health information is supported by a distinct IT management solution. The second triangle in Fig. 1 shows three IT management solutions in the proposed project and they are: 1) the decision-making solution for various illness concerns, 2) the pattern recognition solution for symptoms identification, and 3) the data collection solution for collecting data from ODLs. The personal health information and the IT management solutions are guided by personal health information management strategies as shown in the right portion in Fig. 1. Based on the granularity of the information and the management solutions, these management strategies have three different layers: knowledge management, information management, and data management. Lower-layers of personal health information, IT solutions, and their management strategies are the basis of upper-layers, respectively. For example, the collected ODLs are the basis for identifying the patterns of the symptoms, and the patterns of the symptoms are the basis for making decision regarding patients' illness concerns. Integrated with the management strategies for personal health information and IT solutions mentioned above, our proposed IT management solution for personal health information should be responsive to the ODLs data, the patterns of symptoms, and the illness concerns of the target patients.

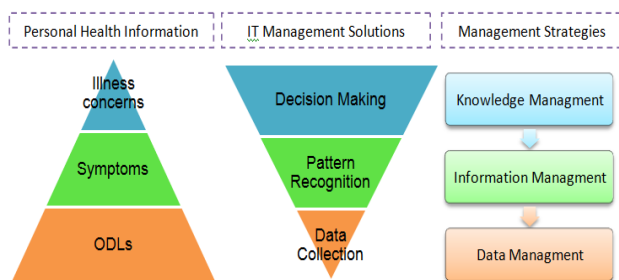


Fig. 1. Personal health information, IT management solutions and their management strategies

As ODLs are the basis for the symptom identification and the decision-making for the illness concerns of geriatric patients with DAH, we must identify a range of ODLs that have the potential to lead to insights for patients, caregivers, and clinicians. In Table 1, the ODLs are listed and categorized in terms of symptoms of DAH [1][2]. These ODLs are meaningful to both the individuals and their professional healthcare providers as they are closely related to the individuals' daily living and also associated with the symptoms of DAH. It is believed that new patterns can be discovered and new knowledge can be obtained from those ODLs.

The use of ODLs has the potential to change clinical practice for our clinical partners:

- 1) Improving the effectiveness of the treatments - Since IT solutions at various levels will be provided for collecting accurate and consistent ODLs data, mining patterns of the symptoms, and making accurate health-related decisions, clinicians will be able to offer far more effective treatments for geriatric patients with DAH.
- 2) Improving the efficiency of the treatments - Clinical visits for patients would not be required as the ODLs are collected in a non-clinical setting. This facilitates remote

diagnosis and remote treatments for patients and the efficiency of the treatments can be greatly improved.

Table 1. Categorized ODLs based on diseases

Diseases	Categories	Symptom Measurement Dimensions	ODLs
Dementia	Cognitive	wandering, falling, cleaning condition patterns	walking, personal hygiene
	Psychiatric	mood change patterns	facial expression
	Functional	sleeping, diet, medication compliance patterns	sleeping, eating, medication taking
Hypertension	Uncontrolled blood pressure	mean / average blood pressure readings	blood pressure
	Inappropriate lifestyle	diet, exercise	caloric intake, calories burned

A. User-centered Requirement Gathering Process

Gathering and understanding the user needs, including the needs of patients, caregivers, and clinicians, is the first step of our technical approach. The user needs determine the direction and overall quality of the resulting system. As shown in Fig. 2, we place the users at the center of the requirements gathering process, and other related elements in the process, such as designers, environments, systems, and interfaces, are influenced by the users. Understanding relationships between users and other elements helps us to answer questions related to requirements, such as 1) how to narrow the gaps between users' expectations and designers' capabilities (i.e., users and designers), 2) in which circumstance, users would like/dislike/hate using the motion sensor devices (i.e., users and environments), 3) how often users use the system (i.e., users and systems), and 4) if the system is user-friendly (i.e., users and interfaces). There are several techniques that can be used for gathering user needs, including interviewing, workshop, questionnaires, brainstorming, storyboards, and screenshots [7].

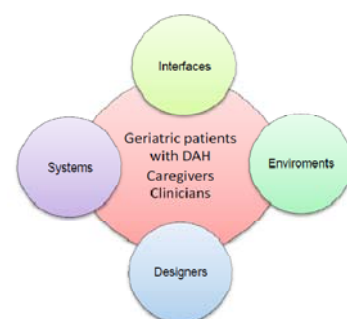


Fig. 2. User-centered requirement gathering strategy

B. Iterative Design Process

The process of gathering and understanding the user needs is inseparable from the design process, since the user needs should be verified and often are required to be adjusted in the design process. To ensure the proposed design is sufficient to

yield new insights to meet the user needs, an iterative design approach is used to incrementally explore the possible ODLs and gain greater understanding of DAH. Our proposed iterative design approach for the system [8], as shown in Fig. 3, starts with user stories. The desired requirements should be produced by the end of the each iteration. The user stories are comprised of short descriptions of user needs and a set of test cases to evaluate the proposed ODLs. Once the user stories have been identified, four activities form the main cycle of work:

- 1) Identify users who will use the product, what they will use it for, and under what conditions they will use it.
- 2) Specify business requirements that must be met for the product to be successful.
- 3) Convert requirements to a complete design, i.e., a release.
- 4) Carry out user-based assessment of the release. This activity is performed by an evaluation component.

The evaluation component consists of a quantified comparison study to measure effectiveness of the ODLs and an empirical study for testing the usability of the system. The evaluation results determine whether those four activities in the cycle need to be performed again for the current release. The whole design process re-starts when new user stories are introduced. The iterative design approach reduces the risks of design failures as it allows for the complete overhaul and rethinking of design by early testing of conceptual models and design ideas. Such an approach greatly facilitates innovative designs to obtain new understanding of DAH.

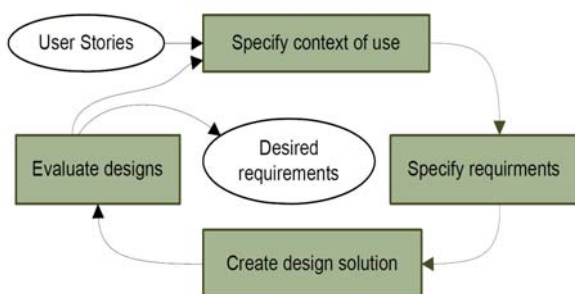


Fig. 3. Iterative design process

C. System Dataflow

The architecture design of the proposed system is critical to the success of the project. Our architecture design starts with understanding the data flow of the proposed system (see Fig. 4).

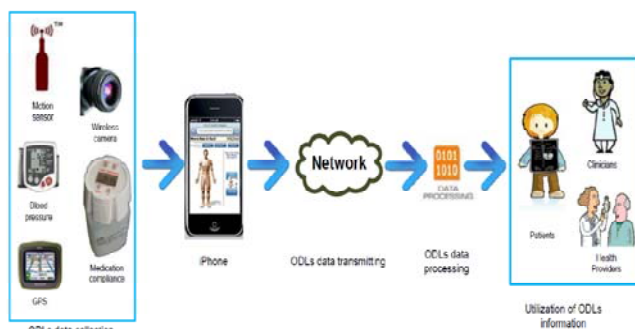


Fig. 4. System dataflow in proposed system

Based on the direction of the system data flow, the proposed architecture design needs to address the following design challenges that are associated with the ODLs transformation: 1) how to use devices, including cameras, global positioning systems, motion sensors, blood pressure devices, and medication compliance devices, to capture ODLs associated with geriatric patients with DAH; 2) how to store the captured ODLs that requires a huge amount of storage space; 3) how to organize and pre-process ODLs from different devices for data interpretation; 4) how to interpret ODLs in relation to DAH symptoms associated with patients' sleeping, walking, and eating patterns; and 5) how to help health providers make health-related decisions based on the interpretation to improve the living quality of geriatric patients with DAH.

D. Service-oriented Architecture and Interpretation Plan

Our design strategy will consider each data transformation as a service that the proposed system needs to provide. Thus, we are able to exploit service-oriented architecture (SOA) to develop a loosely coupled system for addressing these design challenges. SOA provides methods for systems development and integration. As shown in Fig. 5, SOA components are considered as services in the computing clouds. Each service describes its capabilities (i.e., information and behaviors) in web services description language. SOA hides the internal workings from outside intrusion and presents a relatively simple interface to the rest of the system [6]. The utilization of SOA makes full use of existing systems and libraries. It will decrease the costs and implementation time, and increase the quality and reliability of the system.

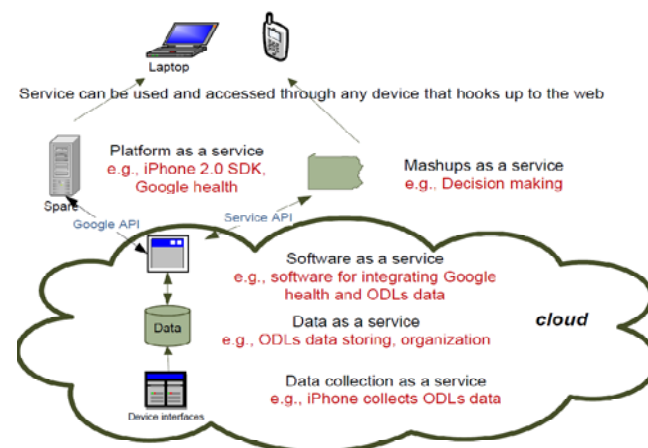


Fig. 5. Service-oriented architecture

We have chosen the iPhone 2.0 software development kit (SDK) and Google Health as the third party platforms for the proposed system (Fig. 5). The iPhone SDK is used as a communication platform to transmit health-related information, such as ODLs, to remote servers. The iPhone SDK provides developers with a rich set of application programming interfaces (APIs) and tools to create innovative data communication applications. Specifically, the iPhone's application APIs and other libraries facilitate us to take patients' facial and food images, record patients' walking and sleeping information with iPhone's embedded motion sensors, locate patients' positions in maps, etc. We also integrate the Google Health platform into the proposed

system. Google Health provides some key features for the proposed project such as: 1) free and secure services to store and manage all of patients' health profiles in one central place, 2) built-in mechanisms for sharing the profile information with other health agencies, and 3) data APIs (.NET and Java) for additional services from other vendors, such as analyzing patients' health histories. These APIs are extremely useful for the proposed system since our ODLs need to be integrated with patients' health profiles. In addition, other Google Data APIs can be used in our project, i.e., Google Map API for tracking positions of geriatric patients with DAH. In our design approach, these features are treated as services so that they can be easily replaced by other similar services, if necessary. The main trade-off of the design is between the use of existing artifacts and customized service. For example, Google Health does not support user-defined data structures, and ODLs cannot be stored in the Google Health center repository. As such, a data integration service needs to be realized within the proposed project for connecting patient health profiles with the ODLs collected.

Our ODL interpretation plan targets on understanding ODLs at different levels (see Fig. 1), including ODL patterns, DAH symptoms, and decision-making strategies. More specifically, our plan answers the questions similar to: 1) how to determine whether the geriatric patient with DAH is sleeping, 2) if the patients have certain sleeping patterns, 3) whether the sleeping patterns are associated with particular DAH symptoms, and 4) what health-related decisions could be made based on the sleeping patterns. The proposed guidelines for the data interpretation will cover the following aspects: 1) Evaluating the preliminary data. A set of criteria should be developed to determine whether the collected ODLs are qualified and sufficient for interpretation purposes; 2) Choosing interpretation techniques. Appropriate methods, algorithms, and tools will be chosen, such as data classification and predication, in terms of the hypothesis of the interpretation; 3) Interpreting data. Applying techniques to interpret the data; 4) Interpreting results. We define contexts and constraints in which the ODLs would lead to the conclusions; 5) Evaluating results and reporting. To evaluate the results, we need to formulate conclusions resulting from comparisons of observation and control groups; and 6) Making health-related decisions. Making correct decisions regarding the health concerns of individual patient is one of our goals. We have adopted the SWOT analysis approach [9] as our health-related decision-strategic approach. The SWOT consists of four elements, including strengths, weaknesses, opportunities, and threats. The SWOT analysis is an extremely useful tool for understanding and decision-making for all sorts of situations in business and organizations. In our data interpretation plan, the strengths and weakness are referred to the health condition of target patients; the opportunities are the possible treatments based on the discovered ODLs patterns and new knowledge derived from ODLs; and the threats are the potential damages and harms learned from the ODLs. The analysis results of each element depend on the IT management solution we have mentioned earlier.

E. Technical Challenges and Solutions

Our overall design strategy is to consider each of the challenging issues as a service that the proposed system needs to provide. This can effectively minimize the risk of system failure as similar services can be easily replaced, if necessary, for capturing and storing ODLs, integrating ODLs with Google Health, and interpreting ODLs. Nevertheless, there are technical challenges associated with these services. As examples, Table 2 shows some of such challenges and possible solutions.

Table 2. Technical challenges and solutions

Challenge Types	Technical Challenges	Possible Solutions
Data capturing	How to capture patients' sleeping data	Using motion sensor embedded iPhones
Data storing	How to store the sleeping data	Storing data in iPhones, and then transmitting it to servers
Data integration	How to integrate ODLs with profiles stored in Google Health	Implementing an integration service, and using Google Health Data APIs
Data Interpretation	How to measure a "good" sleep	Dividing sleep to smaller phases that are measurable

F. Integrating ODLs into Clinical Practice and Risk Control

The workflow of the clinical practice consists of four components: *pre-visit*, *check-in*, *treatment*, and *check-out* (see Fig. 6). To integrate the ODLs into the clinical practice workflow, we first divide the ODL process into four continuous tasks: *data pre-processing*, *analyzing*, *decision making*, and *report generating*. Then, each task is incorporated into one of the appropriate workflow components. The integrated ODLs are the basis for processing ODLs, which is located in the center of the workflow. The integration shown in Fig. 6 is feasible and easy to implement since 1) the components of workflow are well known, 2) each task of the ODL processing is well defined, and 3) each task fits the context of the components of the clinical practice workflow.

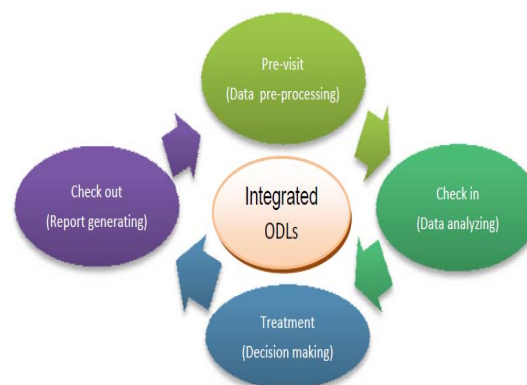


Fig. 6. Integrating ODLs into clinical practice workflow

To reduce risks in implementing ODLs in clinical practice, we first define a risk model for identifying critical risks. The risk model shown in Table 3 describes four types of risks, including *privacy*, *treatment*, *user experience*, and *system performance*. With a well-formulated risk model, the classifications of a risk can be determined and thus, an optimal plan to mitigate the risks can be devised. For example, the utilization of authentication and authorization mechanisms can prevent the unauthorized tracking.

Table 3. Risk model in implementing ODLs in clinical practice

Risks	Risk Descriptions	Proposed Risk Mitigation Plans
Privacy	Compromising patients' daily information or data, such as loss of personal behavioral/mental information and unauthorized tracking	Define different levels of authentication and authorization, generate log files
Treatment	Inaccurate health decisions due to the individual health variants	More examples are required for making accurate decision
User experience	Unwilling learning, inappropriate operating, and insufficient maintenance related issues	Motivation and training should be provided for users
System performance	Unsatisfied performance of the system	More CPU powers/memory and system tuning

We plan to demonstrate the robustness of our technical approach by illustrating how risk control mechanisms are incorporated into the development process in order to cope with both predicable and unpredictable risks. More specifically, we utilize an evolutionary software-process model [11] shown in Fig. 7 to identify, manage, and reduce the technical risks. The evolutionary software process is iterative, and involves repeated cycles of liaison, planning, risk analysis, construction, and evaluation. Each iteration provides a sub-optimized solution for the current release. It prevents risks from propagating into next iteration. Thus, the iterative and incremental development approach ensures the robustness of our technical approach.

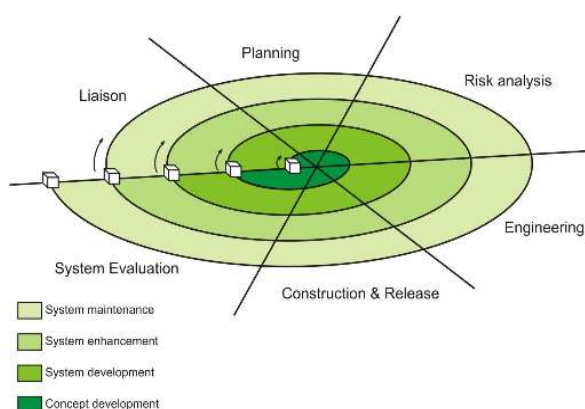


Fig. 7. Evolutionary software process model

Note that, additional to the risks and challenges we have identified above, there are also other challenges relevant to practices and policies associated with the overall technical approach. For example, to collect high quality ODLs, one of the policies in our approach is requiring patients to carry an iPhone all the time; whereas practically, it is difficult for us to check unless we pay extra attention on that, e.g., develop a program to detect if patients carry an iPhone.

IV. EVALUATION DESIGN

The key ingredients of a successful plan for the design evaluation include a sufficient number of participants, a sufficient time frame of ODL collection, ODL control groups, and the equivalent groups at pre-screening. Determining the sufficient number of participants and the sufficient time frame of ODL collection in the experiments are critically important and difficult components of the evaluation design. In general, these key ingredients are described as follows:

- 1) Sufficient number of geriatric patients with DAH. Evaluation designs with small numbers of participants often have inadequate statistical evidence to detect significant changes across ODL observation groups and control groups. The ODLs analysis results obtained from the geriatric patients with DAH in our research group should be reflected to and readily generalize to other population;
- 2) Sufficient time frame of ODL collection. Similar to the reasons for requiring the sufficient number of participants in the project, evaluation designs with short time frames often have inadequate statistical evidence to detect significant changes over time. In our proposed ODLs, some of abnormal behaviors/patterns do not occur very often or are not even visible for some patients, e.g., getting lost while walking. It takes time to collect walking data that may contain wandering patterns.
- 3) ODL control groups. The medication compliance monitoring, blood pressure pattern monitoring, sleeping pattern monitoring, and other prevention and intervention treatment typically have slow effects on patients with chronic illnesses, as do the related risk factors. The best way to evaluate prevention and intervention effects is to compare changes over time in those who receive prevention and intervention treatments (i.e., receive feedback and the ODL-related health decisions from healthcare providers) with those in a control group who do not receive such treatments. Within a control group design, it is our expectation that risk factors will be reduced in the prevention and intervention group compared to the control group. In other words, the proposed project will be effective in preventing and reducing the injury in target patients in a particular problem behavior.
- 4) Equivalent groups at pre-screen. The prevention and intervention and control groups should be equivalent in size, demographics, and disease symptoms, as much as possible before the intervention starts. Demonstrating pre-screen equivalence between two groups helps to establish that any differences in ODLs, ODL pattern difference, health decision differences, and treatment effectiveness difference after the prevention and intervention are due to the prevention and prevention program and not to pre-screen differences.

Sample size determination is an important issue in our proposed project. First, it must be "large enough" that an effect is of such magnitude as to be of scientific significance will also be statistically significant. On the other hand, it

cannot be “too large,” where an effect of little scientific importance is nevertheless statistically detectable [11]. From the economic perspective, an under-sized study can be a waste of resources for not having the capability to produce useful results, while an over-sized one uses more resources than are necessary.

There are several approaches to determine the sample size, including a confidence interval approach that the sample size is determined by the desired width of a confidence interval and a Bayesian approach to predicate the sample size. One of the most popular approaches to sample-size determination is based on the power of a test of hypothesis [12][13].

The determination of sufficient time frame of ODL collection is similar to determine the sample size. Besides the step we list above, we need to estimate the frequency of the symptom measurement dimensions (see Table 4) in ODLs, such as falling, wandering, uncontrolled blood pressures, and other inappropriate lifestyle. The duration of ODL collection is based on the required sample size and the frequency of the symptom measurement dimensions in ODLs.

The number of participants and the time frame of ODLs collection are also determined by data analysis techniques. Data analysis is a process of gathering, modeling, and transforming data with the goal of highlighting useful information, suggesting conclusions, and supporting decision-making. Techniques of data analysis range from simple measures, such as means and standard deviations, to more complex analyses such as regression. Data mining is a particular data analysis technique that focuses on modeling and knowledge discovery for predictive rather than purely descriptive purposes [14]. Theoretically, the accuracy of the prediction highly depends on the amount of the data available. Data mining commonly involves four classes of task. We describe each task and possible applicable ODL symptom measurement examples in Table 4.

Table 4. Four classes of data mining task and their applicable ODL symptom measurement examples

Tasks	Description	Commonly Used Algorithms	Symptom Measurement Examples
Classification	Arrange the data into predefined groups nearest neighbor	Naïve Bayes classifier	Sleeping pattern
Clustering	Classify without predefinition	Neural network	Falling and diet pattern
Regression	Attempt to find a function, which models the data with the least error	Genetic programming	Mood change pattern
Association rule learning	Search relationships between variables	Apriori algorithm	Medication compliance vs. cleaning condition

Our evaluation plan is reliable and easily carried out because we have addressed the following issues: 1) Sufficient participants. Our clinical partner is able to provide sufficient patients from which to choose our target patient population; 2) Sufficient time frame. Overall, a six-month experiment period should provide sufficient number of ODLs for data analysis; 3) Convenient data collection devices. All data collection devices are small enough for patients to carry; 4)

Fewer burdens for patients. All the devices proactively collect ODLs, and geriatric patients with DAH need to do almost nothing during the experiments; 5) Reuse of data mining tools. Publicly available data mining tools can be used for our project, such DataCruncher, WEKA, OLPARS, and S-Plus.

V. CONCLUSION

This paper demonstrates a novel treatment approach to provide cost-effective quality therapies for geriatric patients with DAH. The goal of the proposed project is to seek to exploit current information technologies, including theories, algorithms, processes and tools, to develop a system for testing whether and how ODLs collected from a non-clinical environment can be interpreted and utilized to help patients with DAH remain in their homes safely and improve their life quality while reducing medical expenditures. A service-oriented architecture design is employed to make full use of existing devices, software systems, and platforms, including the utilization of the iPhone 2.0 software development kit for the ODL collection and transmission and the Google Health platform for storing personal health profiles. The outcomes from the proposed work will have a great potential for commercialization, since the existence of market is huge and this health care system is cost effective.

REFERENCES

- [1] D. M. Eppele, “Senile dementia of the Alzheimer type,” *Clin. Soc. Work J.*, vol. 30, pp. 95-110, 2004.
- [2] K. Maslow, “Dementia and serious coexisting medical conditions: a double whammy,” *Nurs. Clin. North Am.*, vol. 39, pp. 561-579, 2004.
- [3] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth edition. Washington, DC: American Psychiatric Association; 1994.
- [4] E. Farina, R. Fioravanti, R. Pignatti, M. Alberoni, F. Mantovani, G. Manzoni, L. Chiavari, E. Imbornone, F. Villanelli, and R. Nemni, “Functional living skills assessment: a standardized measure of high-order activities of daily living in patients with dementia,” *Eur J Phys Rehabil Med.*, vol. 46, 2010.
- [5] Robert Wood Johnson Foundation. “Feature: The Power and Potential of Personal Health Records.” January 2010. [Online] available at <http://www.rwjf.org/pr/product.jsp?id=51248>
- [6] T. Erl, *Service-oriented architecture: concepts, technology, and design*. Upper Saddle River, NJ: Prentice Hal, 2005.
- [7] M. Maguire and N. Bevan, “User requirements analysis. A review of supporting methods,” *Proceedings of IFIP 17th World Computer Congress*, Montreal, Canada, August 2002, Norwell, MA: Kluwer, pp.133-148, 2002.
- [8] J. Nielsen, Iterative user interface design. *IEEE Comput.* 26(11), pp. 32-41, Nov, 1993
- [9] R.P. Mishra, G. Anand, and R. Kodali, “Strengths, weaknesses, opportunities, and threats analysis for frameworks of world-class maintenance,” *Proceedings of the Institution of Mechanical Engineers – Part B – Engineering Manufacture*, 221(7), pp. 1198-1208, 2007.
- [10] B. Boehm and W. Hansen, *The spiral model as a tool for evolutionary acquisition*, CrossTalk, 2001.
- [11] D. J. DeSousa, A. A. Rouse, and W. J. Smolon, “Statistical consequences of reducing the number of rabbits utilized in eye irritation testing: Data on 67 petrochemicals,” *Toxicol. Appl. Pharmacol.* vol. 76, pp. 234-242, 1984.
- [12] R. E. Odeh and M. Fox, *Sample size choice*. New York: Marcel Dekker, 1991.
- [13] J. A. Freiman, T. C. Chalmers, Jr. H. Smith, and R. R. Kuebler, “The importance of Beta, theType II error, and sample size in the design and interpretation of the randomized controlled trial: Survey of 71 “negative” trials,” in *Medical Uses of Statistics*, eds. J. C. Bailar III and F. Mosteller, NEJM Books, Waltham, Mass. chap. 14, pp. 289-304, 1978.
- [14] U. Fayyad and R. Uthurusamy, “Data mining and knowledge discovery in databases,” *Comm. ACM.*, vol. 39, pp. 24-26, 1996.