A Model-Based Systems Engineering Framework for Healthcare Management with Application to Diabetes Mellitus

Iakovos M. Katsipis and John S. Baras
2247 A.V. Williams Building
Institute for Systems Research
University of Maryland
College Park, MD 20742, USA
Tel: +1-301-405-6606; Fax: =1-301-314-9218) – email: ikatsipi@isr.umd.edu; baras@umd.edu

Abstract: We describe a framework for Health Care Management Systems using modern Model-Based Systems Engineering methodologies and apply it to Diabetes Mellitus. We describe the desired architecture of such systems. We include a Controlled Hidden Markov Chain model for diabetes disease progression with three states, three diagnostic tests, ten interventions, three types of patients. We develop three methods for computing tradeoffs between health care cost and health care quality. One is an exhaustive Monte Carlo simulation and the other two use multi-criteria optimization with full and partial disease state information. The latter obtain similar results at a fraction of the time of the first. We demonstrate the powerful capabilities of such a framework via examples and problems of practical healthcare significance.

Keywords: model-based systems engineering**,** decision making**,** trade-offs, disease progression models**,** health care quality metrics, cost metrics

1. INTRODUCTION AND SIGNIFICANCE OF THE PROBLEM

Improving health care delivery and management is a major goal of governments around the world, and a centerpiece of many reform efforts [1-7]. Healthcare systems around the world are facing unprecedented challenges and global issues [2-5]. Healthcare costs are rapidly increasing (rose 2.6% in 2013, accelerating to an average of 5.3% per year over 2014-2017) as years go by, and unfortunately coverage and offered services are decreasing. There are four major issues that governments, health care providers, payers, and consumers face: aging population and chronic diseases; cost and quality; access to care; and technology. A key trend creating increased health care demand is the spread of **chronic diseases** [2]: heart disease, stroke, cancer, chronic respiratory diseases, **diabetes**, and mental illness, among other. This trend has major implications on healthcare costs.

Health care technology changes will be rapid and, in some parts of the world, disruptive to established health care models [1-7]. Some exciting advancements are taking place at the intersection of information technology and medical technology. Providers can leverage vast amounts of patient data, gathered from a variety of sources, to determine the clinical value of specific treatments and how to improve them [1-7]. Adoption of new digital health information technologies (HIT) such as electronic medical records (EMRs), telemedicine, mobile health (mHealth) applications, and electronic medical prescriptions is driving change in the way physicians, payers, patients, other stakeholders interact [1-7]. **Health Information Technology** (Health IT) [1-7] has great potential to ameliorate these problems, and is being aggressively pursued in the US, Europe and other countries [1-7]. Health IT is of central interest for the problems addressed in this paper. Health IT systems are complex systems and even systems of systems [1] and need to be treated as such; that is taking a **holistic and integrative systems view** [1]. The challenge is even greater because humans of various capabilities, functionalities and roles are essential parts of health care systems.

This grand challenge provides the main motivation for the work reported in this paper. We address a specific class of health care management systems (HCMS), as a first but important step towards the systematic application of modern **Model-Based System Engineering** (**MBSE**) methodologies, frameworks and tools for the design, construction, operation and maintenance of such systems [1]. We selected as a focus application the modeling and management of **Diabetes Mellitus** (or **Diabetes 2**) for its high impact, as it affects tens of millions of people world-wide. In the USA alone [8], 29.1 M people (or 9.3% of the population) have diabetes, with estimated annual diabetes cost of $245B (2012 data). What is even more important is the fact that it is predicted that both figures will keep growing dramatically [9, 10].

We describe a methodology and a framework that utilizes recent advances in MBSE and associated tools, to develop a conceptual architecture for such a HCMS for Diabetes 2 with the following capabilities and characteristics:

(i) Is scalable to millions of patients, and tens of thousands of healthcare providers;
(ii) Is expandable, in the sense that it can continuously accommodate new data and knowledge, new tests, new, models, new treatments;
(iii) Is linkable to distributed medical databases;
(iv) It has capabilities to "learn";
(v) It can be easily used by healthcare providers, health insurance managers, patients;
(vi) It can operate in a distributed collaborative manner and be linked to extensive communication and data networks and large heterogeneous sensors and databases;
(vii) It can provide quantitative answers to "what-if" type of questions such as: what is the effect of using modern monitoring wearable technology, what is the most effective test, what is the most effective treatment, what are the tradeoffs between costs and tests and treatments.

We further focus on a key component of such a HCMS, which is the **Reasoning Engine** to perform efficiently the required difficult tradeoffs in many key decisions. We want to demonstrate the huge potential of systematically utilizing modern MBSE methods towards reaching the goals and needs described in [1], and encourage the use and development of such systems linked to real-life data over extensive periods of time. The overall architecture of the type of HCMS we have in mind, as illustrated in Figure 1, provides a variety of connectivity to facilities, labs, hospitals, shareholders (patients, doctors, insurance managers, etc.). The dotted lines indicate a modern communication and data network between healthcare actors. Such networks have been developed and are under development in many countries including the USA, as an example in Massachusetts [11-13].

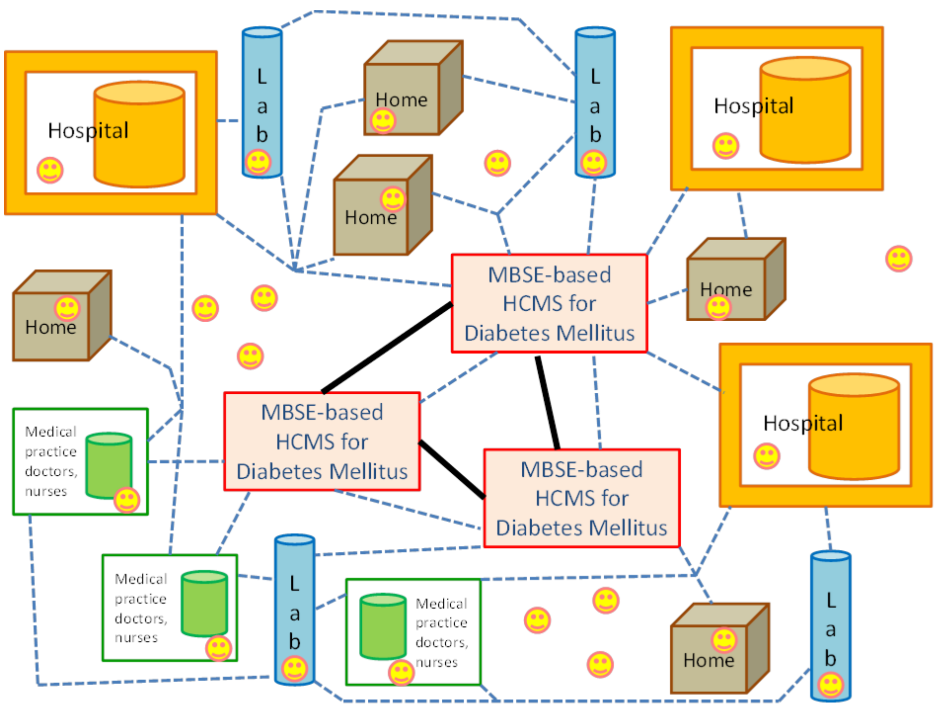

Fig 1: Illustrating the MBSE-based Health Care Management System for diabetes mellitus and its functional connectivity.

The organization of the rest of the paper is as follows. In Section 1, we describe the MBSE approach we follow, in Section 2 we describe the disease model, the metrics we used, and trade off analysis methods we developed and used, in Section 3 we provide some representative results of tradeoff analysis.

1. APPROACH: MODEL-BASED SYSTEMS ENGINEERING AND SYSTEM ARCHITECTURE MODEL AS A SYSTEM INTEGRATION FRAMEWORK

The term “Model Based System Engineering” could be described as a rigorous, quantitative representation of system structure and behavior components to support system requirements management, design, verification and validation activities beginning with the conceptual design phase and continuing through-out development, operations, and later life cycle phases [14-17]. In comparison to the document centric approach, here the component models and their interconnections are the main artifact of each procedure and are used for the communication between the dissimilar groups that participate in the development. In Figure 2 the essential phases of the MBSE process are indicated [14-17]. For each system the initial step of the process is the offered evidence. Subsequently, the initial system requirements and the anticipated measures of effectiveness (MoE) are developed (captured). The MoE are used subsequently at the trade-off phase to guide the selection of components, selection of design parameters and other design space exploration functions.

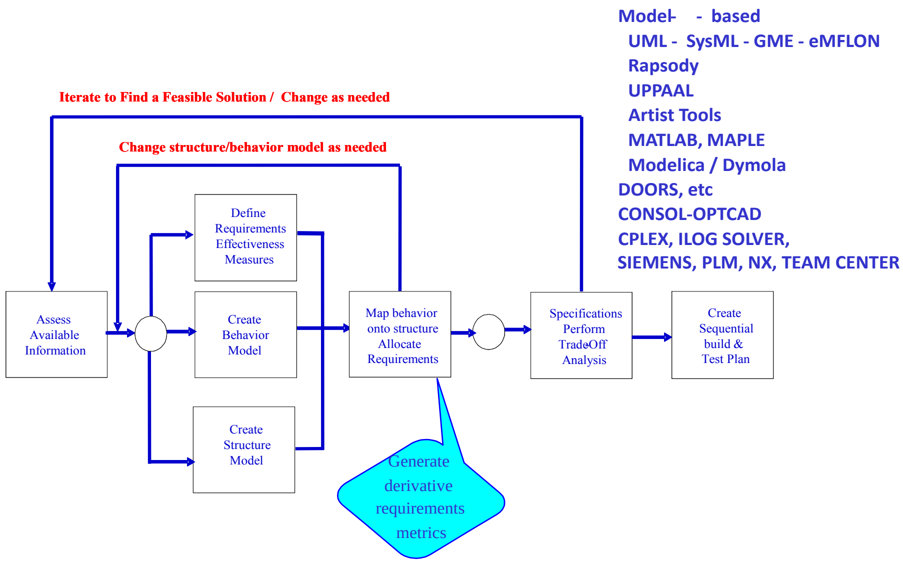

Fig. 2: MBSE process information centric abstractions

System architecture is the abstract model that describes the structure, behavior, their components and interconnections, the mapping of behavior onto structure and various views of a system. An architecture narrative is a formal explanation and illustration of a system, organized in a way that supports rationale about the structure and behavior of the system. System architecture can include system components, the outwardly perceptible properties of those components, the relationships (e.g. the behavior) between them.

Figure 3 shows the essential components of the System Architecture integration framework including hardware models, software models, analysis models, verification, and requirements components. The main challenge and need is to develop scalable holistic methods, models and tools for enterprise level system engineering. Therefore integration of multiple domain modeling tools, trade off tools, validation/verification tools, databases and libraries annotated and component models from all disciplines is required.

2. EVALUATION METRICS, PERFORMANCE EVALUATION AND TRADEOFF ANALYSIS

A key component of our framework is the modeling of disease evolution incorporating medical tests and interventions. We have incorporated the advantages and relevant information from many models and we have integrated them into a new system model that is more complete and detailed [18-26]. We have developed a model for diabetes progression as a Controlled Hidden Markov Chain (CHMC) [27]; see Figure 4. The model has three states for diabetes mellitus, The states we use have the following interpretation: ***State 1*** represents the **Healthy** (disease free) condition of a generic patient. ***State 2*** represents the **Pre-diabetic** condition of a generic patient. ***State 3*** represents the **Diabetic** condition of a generic patient.

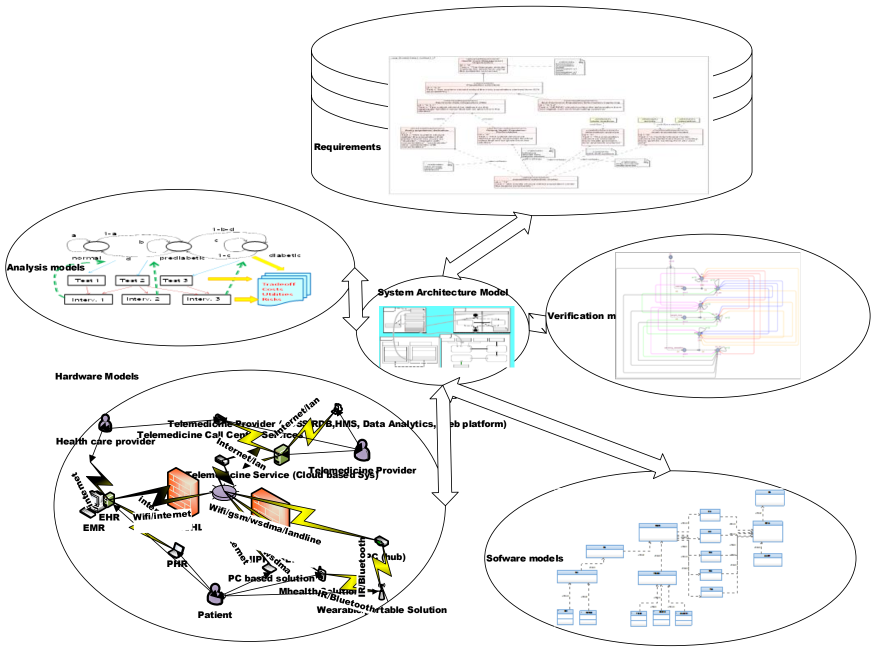

Fig. 3: System architecture integration framework for a healthcare management system

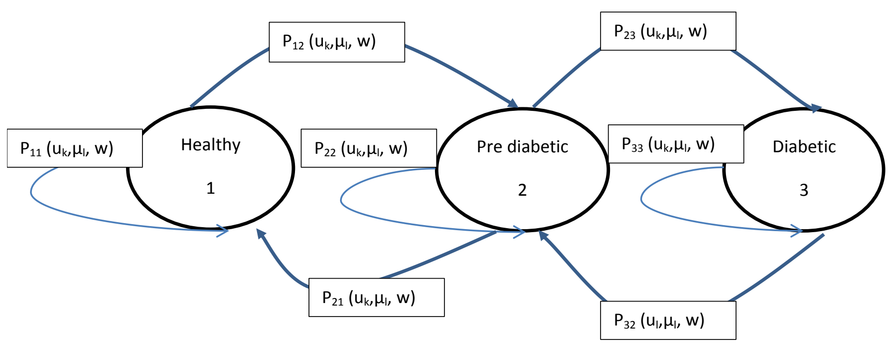

Fig. 4: Graphical representation of the CHMC model developed for Diabetes 2 progression

Our framework provides a systematic methodology for evaluating the quality of health care provided to a generic patient (meaning the sequence of tests and interventions applied), the associated costs, associated tradeoffs and many other important evaluations of individual tests and interventions, over a particular time period. The purpose of our study is not to provide a specific medical model with all its complexities and ramifications, but rather develop, use and demonstrate the significant benefits of a MBSE framework for analyzing these problems. Furthermore, our MBSE framework is by construction **modular** and **composable** allowing many tools and models to be easily integrated within its structure. This is in great contrast with the current state of the art in healthcare IT systems, even with the most advanced ones.

Our analysis specifies a time horizon for the study denoted by *T*; in most of our experiments we take *T* to be ten years. It is defined by the user of the framework: heal care provider, policy maker.

If the disease state is known, or, as is more appropriate for the most realistic models, estimated on the basis of diagnostic tests results, medical practice standards and the reasoning and experience of medical practitioners result in a ***selection process for these interventions***; one can think of these selections as ranking the interventions in some order of preference.

We have included several metrics in our studies; our MBSE framework is particularly suitable for multiple metrics. One performance metric that we are interested in, is **cost**. The cost model we use is pretty straightforward. There is a cost for every test and every intervention used, denoted by $C_u(u_k) = c_{u_k}$ and $C_\mu(\mu_l) = c_{\mu_l}$ respectively. The cost model we use is additive, that is the total cost for anyone time history for each patient is the sum of the cost of the tests, denoted by $C_\mu^{total}(i,m_i)$, and of the cost of the interventions, denoted by $C_u^{total}(i,m_i)$, used in the particular time history m_i. Our cost values include equipment, personnel and other costs. Clearly one could easily develop more detailed models by breaking down the components of each cost. Also, we can use costs data from doctors and hospitals directly in our model, or indirectly by using them to fit a functional model of each cost and its components. Thus for our work and analysis we have

$C_\mu^{total}(i,m_i) = \sum_{t=1}^{N_{T,\Delta}} C_u(u(t))$, $C_\mu^{total}(i,m_i) = \sum_{t=1}^{N_{T,\Delta}} C_\mu(\mu(t))$, and $C^{total}(i,m_i) = C_\mu^{total}(i,m_i) + C_u^{total}(i,m_i)$

Where, in these sums, the tests and interventions used at each time step of this particular time history are considered.

Patients are of different types. Indeed behavioral characteristics of patients with respect to their healthcare are a very important, albeit very difficult to model, factor. In this paper we consider **three types of patients (**the **"Risk Averse"** (with respect to the risk of getting sick with diabetes) patient, the **"Risk Indifferent"** patient and the **"Risk Taker" patient)** with respect to the attention and systematic care that they apply to their health care. We introduce weights representing the value (or significance) each patient places for being in each state of the model (recall Healthy, Pre-diabetic, Diabetic) V_1^i, V_2^i, V_3^i , respectively. These weights take real nonnegative values between 0 and 1, and their values sum to 1. They are a simple, albeit efficient representation of the "profile" of each user.

The number of periods that each patient, in each generated time history, finds herself/himself in each of the three states is an important health condition metric. As state transitions depend explicitly on the tests and interventions applied at each time period, the three counting statistics below, for each patient and each time history, constitute a practical and useful health care value (or quality) metric.

$$O_1^i(m_i) = \text{number of periods, from } N_{T,\Delta} \text{ total, patient } i \text{ is at state 1 (i.e. is Healthy)}$$
$$O_2^i(m_i) = \text{number of periods, from } N_{T,\Delta} \text{ total, patient } i \text{ is at state 2 (i.e. is Pre-diabetic)}$$
$$O_3^i(m_i) = \text{number of periods, from } N_{T,\Delta} \text{ total, patient } i \text{ is at state 3 (i.e. is Diabetic)}$$

Using the weights V_1^i, V_2^i, V_3^i and these counting statistics O_1^i, O_2^i, O_3^i, we can define several metrics for health care quality value for each patient and each time history, generated by our model. For example we consider the following ***Health Care Quality metric***:

$$J_{hc}(i,m_i) = V_1^i * O_1^i(m_i) + V_2^i * O_2^i(m_i) + V_3^i * O_3^i(m_i)$$

We developed and exercised three methods for tradeoff analysis. One of the methods uses the model in an exhaustive straightforward generation of all possible sample paths (time histories) for any number of patients, which we call Evaluation by Monte Carlo simulation (EMCS). The EMCS method has the following steps: ***Step 1***: Run the model for the number of patients, horizon, time step, set of tests, set of interventions, and transition probabilities provided. ***Step 2***: Store the results of Step 1, as triples of arrays (vectors). The first contains the sequence of health states of this specific patient and the specific time history, the second the sequence of tests used at each time step and the third the sequence of interventions used at each time step. ***Step 3***: Using these arrays compute the cost $C_\mu^{total}(i,m_i)$, $C_u^{total}(i,m_i)$, $C^{total}(i,m_i)$, the counting statistics $O_1^i(m_i), O_2^i(m_i), O_3^i(m_i)$, and the healthcare value metric $J_{hc}(i,m_i)$. ***Step 4***: Plot for each patient and time pair (i,m_i), the pair of values $(C^{total}(i,m_i), J_{hc}(i,m_i))$ in the positive quadrant of the plane (where the vertical axis (*y*-axis) corresponds to the total cost C^{total} and the horizontal axis (x-axis) to the Health Care Quality metric J_{hc}) and determine the Pareto points. Figure 5 provides a graphical illustration of the EMCS method.

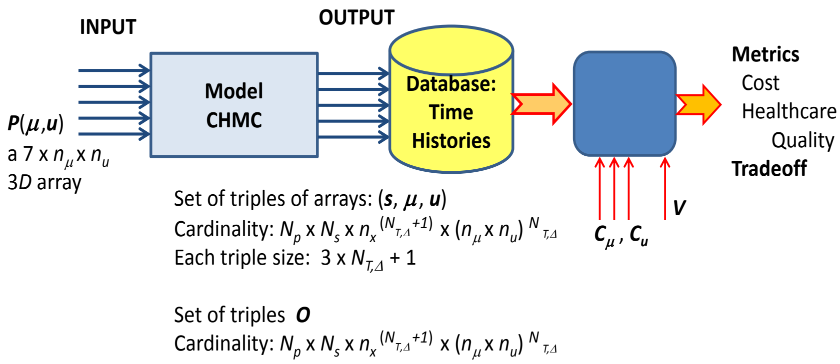

Fig. 5: A graphical illustration of the EMCS information processing flow

The number of patients we have used is 10,000 or 100,000, the time period is 10 years. We have 3 tests that can be given to a patient (denoted by μ), we have 10 interventions that can be applied to a patient (denoted by u), and 3 disease states (denoted by x). A set of 30 matrices describing state transition probabilities, parametrized by the tests and interventions is an input. The second important input is the values for the costs for each test available and every intervention available. The third important input is input of variables used for the computation of the Health Quality metric. We only need to enter the values for the weights for each patient type.

We have created a very versatile system, which can be used for various types of studies, tradeoffs and answering "what-if" type of questions, as we discuss in the next section.

3. DISCUSSION OF DECISION MAKING AND ANALYTICS CAPABILITIES

A key output form our Reasoning Engine system, employing the ECMS method is the computation of Pareto points that describe succinctly the relative value of treatments and tests vs the overall health care quality of a patients time history. Running ECMC with two metrics (and 2-D graphs) for 10,000 patients and 32 runs, took for the whole experiment 783sec. Running ECMS with three metrics (and 3-D graphs) for 100,000 patients took for the whole experiment 1,385 sec. So EMCS is definitely scalable and much better performance can be obtained with more powerful machines. Nevertheless as we are interested in superior efficiency, we developed and analyzed two alternative methods that **decrease the execution time by two or more orders of magnitude**. The other two methods employ a multi-criteria optimization approach to directly compute the Pareto points and associated selection of tests and interventions. The methods use feedback between the disease states, or the estimates of the disease states respectively, and the selection of the tests and interventions to be applied at each time instance. More precisely the second method automatically computes the tests and interventions to be applied at time t as functions of $x(t)$, and uses explicitly the state of the disease. It is thus called **Fully Observable**. The third method employs multi-criteria optimization also with the important difference that the disease state is not available, and only estimates of the state based on the scores and results of diagnostic tests are available. We show that the second method, which we call **Fully Observable Multi-criteria Optimization (FOMCO)** saves tremendously in computational time in achieving similar tradeoff analysis as the exhaustive simulation based EMCS method.

We are interested in analyzing the tradeoff between different pairs of sequences $(\boldsymbol{m},\boldsymbol{u})$ from the perspective of the total Cost $C^{total}(i,\boldsymbol{m},\boldsymbol{u})$ and average Health Care Quality $\bar{J}_{hc}(i,\boldsymbol{m},\boldsymbol{u})$ metrics. This tradeoff analysis is important as it is necessary in order to find the "best" tradeoffs between the two conflicting objectives of maximizing $\bar{J}_{hc}(i,\boldsymbol{m},\boldsymbol{u})$ and minimizing $C^{total}(i,\boldsymbol{m},\boldsymbol{u})$. These "best" tradeoff points are the **Pareto points** (also known as **non-dominated solutions**). To compute these Pareto points for tradeoff analysis using the FOMCO, we utilized the fastest technique of combining the two metrics in a convex combination, resulting in a single criterion stochastic optimization problem; fully observed. Such methods are called **scalarization methods.** We used Dynamic Programming analytics for both of these methods and computations [28].

The Figures 6 and 7 below show typical tradeoff results from our system using the EMCS method and the FOMCO method respectively.

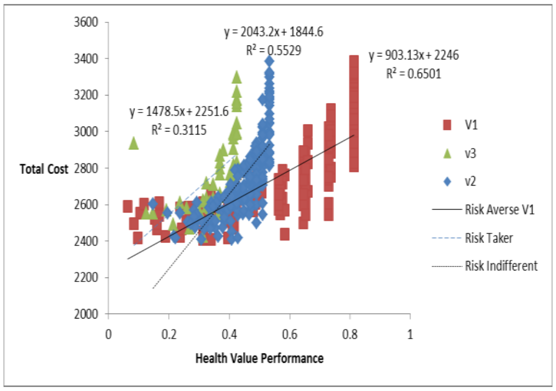

Fig. 6: Cost vs healthcare quality tradeoffs

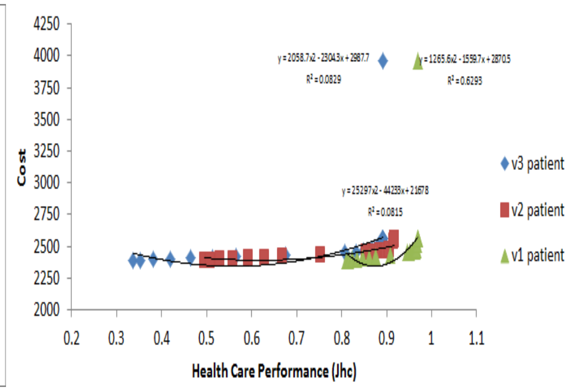

Fig. 7: Cost vs healthcare quality tradeoffs

Using our Reasoning Engine we can provide answers to many practical questions, queries, problems, from health care management perspective, like: Evaluate patient risk behavior impact on health care quality; Evaluate "best" health care achievable; Can learn from new data, treatment results, improve models; Evaluate "value" of new proposed tests and interventions; Provide aggregate statistics for insurance policies calibration; Find best tests and interventions for patient type, disease state; Evaluate effects of incentives and rewards for health "maintenance"; Evaluate sequences of tests and treatments for reversing disease.

Acknowledgment

Research supported in part by NSF under grant CNS-1035655, by NIST under contract 70NANB11H148.

REFERENCES

[1] C. Cassel et al, "Better Health care and Lower Costs: Accelerating Improvement Through Systems Engineering," *Report to the President*, *Executive Office of the President*, *President's Council of Advisors on Science and Technology*, May 2014.

[2] Deloitte, "2014 Global Health care Outlook: Shared Challenges, Shared Opportunities," *Deloitte Development LLC*, *Report*, 2014.

[3] A. Shah, "Health Care Around the World," Global Issues, Article 774, September 2011, http://www.globalissues.org/article/774/health-care-around-the-world .

[4] Z. Jakab, "Facing Future Challenges for Global Health," *World Health Organization (WHO) R&D presentation*, London, *Nov.* 2012.

[5] Intrahealth, "Top 10 Global Health Issues to Watch in 2014," *Intrahealth Organization Report*, January 2014.

[6] KPMG Global Healthcare, "Staying Power, Success Stories in Global Healthcare," *KPMG International*, *Report*, 2014.

[7] KPMG, Healthcare, http://www.kpmg.com/global/en/industry/healthcare/pages/default.aspx.

[8] "National Diabetes Statistics Report, 2014," Division of Diabetes Translation, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention; http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf

[9] American Diabetes Association, http://www.diabetes.org/diabetes-basics/?loc=db-slabnav

[10] CDC Centers for Disease Control and Prevention – Diabetes Home, http://www.cdc.gov/diabetes/home/index.html

[11] The Mass HIway Home, http://www.masshiway.net/HPP/index.htm

[12] The Massacghusetts health Information Highway (The HIway), http://www.mass.gov/eohhs/gov/commissions-and-initiatives/masshiway/

[13] *The Massachusetts Health Information Highway Fact Sheet*, http://mehi.masstech.org/sites/mehi/files/documents/Mass-HIway-Fact-Sheet.pdf

[14] J.S. Baras, Lecture Notes for MSSE class, ENSE 621, 2002.

[15] B. Wang and J. S. Baras, "Integrated Modeling and Simulation Framework for Wireless Sensor Networks", *Proceedings of the 21st IEEE International Conference on Collaboration Technologies and Infrastructures (WETICE 2012- CoMetS track)*, pp. 268-273, Toulouse, France, June 25 - 27, 2012.

[16] D. Spyropoulos and J. S. Baras, "Extending Design Capabilities of SysML with Trade-off Analysis: Electrical Microgrid Case Study", *Proceedings of the Conference on Systems Engineering Research (CSER'13)*, pp. 108-117, Atlanta, GA, March 19-22, 2013.

[17] Y. Zhou and J.S. Baras, "CPS Modeling Integration Hub and Design Space Exploration with Application to Microrobotics," Chapter in the volume *Control of Cyber-Physical Systems*, D.C. Tarraf (ed.), Lecture Notes in Control and Information Sciences 449, pp. 23-42, Springer 2013.

[18] A.I. Adler AI, R.J. Stevens, S.E. Manley, et al, "Development and Progression of Nephropathy in Type 2 Diabetes: the United Kingdom Prospective Diabetes Study (UKPDS 64)," *Kidney Int*63:225–232, 2003

[19] I. M. Stratton, E. M. Kohner, S. J. Aldington, et al, "UKPDS 50: Risk Factors for Incidence and Progression of Retinopathy in Type II Diabetes Over 6 Years From Diagnosis," *Diabetologia*44:156–163, 2001

[20] E. Müller, S. Maxion-Bergemann, et al, "EAGLE Diabetes Model: Basic Features and Internal Validation of Simulating Long-term Diabetic Outcomes and Related Costs," *Value Health*7:745, 2004

[21] E. Müller, et al, "EAGLE–Economic Assessment of Glycemic Control and Long-term Effects: a Computer Simulation Model for Diabetes Mellitus Type 1 and Type 2, *Diabetologia* 47:A355, 2004

[22] E. Müller, et al, "Development and Validation of the Economic Assessment of Glycemic Control and Long-Term Effects of Diabetes (EAGLE) Model," *Diabetes Technol Ther* 8:219–236,2006

[23] A. J. Hayes, J. Leal, A. M. Gray, R. R. Holman, P. M. Clarke, "UKPDS Outcomes Model 2: A New Version of a Model to Simulate Lifetime Health Outcomes of Patients With Type 2 Diabetes Mellitus Using Data From the 30 Year United Kingdom Prospective Diabetes Study: UKPDS 82. ," *Diabetologia*, 2013 Sep;56(9):1925-33. doi: 10.1007/s00125-013-2940-y. Epub 2013 Jun 22. PubMed PMID: 23793713

[24] P. Zhang, X. Zhang, J. Brown, D. Vistisen, et al., "Global Healthcare Expenditure on Diabetes for 2010 and 2030.," *Diabetes Res Clin Pract.* 2010 Mar; 87(3):293-301. doi: 10.1016/j.diabres.2010.01.026. Epub 2010 Feb 19. Erratum in: *Diabetes Res Clin Pract.* 2011 May;92(2):301. PubMed PMID: 20171754.

[25] M. I. Schmidt, B. B. Duncan, H. Bang, J. S. Pankow, C. M. Ballantyne, et al, "Atherosclerosis Risk in Communities Investigators. Identifying Individuals at High Risk for Diabetes: The Atherosclerosis Risk in Communities Study., *Diabetes Care*. 2005 Aug;28(8):2013-8. PubMed PMID: 16043747.

[26] D. Roz, M. D. Lasker, *N. Engl J Med* 1993; 329:1035-1036 September 30, 1993, and *The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study Research Group*. *N Engl J Med* 2005; 353:2643-2653 December 22, 2005.

[27] M. Sordo, T. Hongsermeier, V. Kashyap, and R.A. Greenes, "A State-Based Model for Management of Type II Diabetes," presented at *Advanced Computational Intelligence Paradigms in Healthcare* (1), 2007, pp.27-61.

[28] D. Bertsekas, *Dynamic Programming and Optimal Control*, *Volume 2*, Athena, MA, 1995.