VALIDATING A COMPUTATIONAL MODEL OF PATIENT ILLNESS: THE SIMCARE PATIENT MODEL

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Ryan M. McCabe

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Paul E. Johnson

Paul R. Schrater

July 2012

Acknowledgements

I would like to acknowledge the combined impact my thesis committee has had on my development throughout school and the completion of this degree, especially my two advisors. Paul Schrater has broadened and deepened my knowledge of probability theory, and Paul Johnson has taught me how to develop principled research and write more effectively. Both have made invaluable contributions to my understanding of modeling and multi-disciplinary work.

I offer broad thanks to the many faculty and students involved with the AIRVL lab and Outobox group, especially Maria Gini, Dan Boley, Paul Schrater, and Steve Jensen. Gedas Adomavicius has helped refine my knowledge of and approach to applying various data mining algorithms to real-world problems. I would also like to thank several individuals who have collaborated on much work and countless conversations, including Gregory Ramsey, Georg Meyer, and Mohamed Elidrisi. Drs. Patrick J. O'Connor, Joann Sperl-Hillen and William Rush have contributed greatly to my understanding of patients living with type 2 diabetes. Dr. George Biltz contributed greatly to the conceptualization and implementation of the original SimCare prototype as well as its reimplementation, which served as the version used in this thesis. Thanks to Autri Dutta for correspondence regarding the model.

Dedication

This dissertation is dedicated to my parents, Dennis and Jane McCabe, and my wife, Vanessa. Your personal support on hand and from afar has made the completion of this work possible.

Abstract

The SimCare Patient Model is a computational model of individuals with type 2 diabetes. The model represents a patient as a sequence of health states that respond to treatments over varying intervals of time. It was originally constructed as a "clinical" model of an "individual patient" with type 2 diabetes so that a physician could access the model by querying the patient state for information, ordering specific treatments for the simulated patient and scheduling the next clinical encounter. A software implementation of the model, generated by previous research (Dutta, et al. 2005), has been used as a training tool for medical residents and primary care physicians (O'Connor, Sperl-Hillen, et al., Simulated Physician Learning Intervention to Improve Safety and Quality of Diabetes Care: A Randomized Trial 2009), a guideline and protocol simulator as well as a tool for identifying optimal treatments under given constraints (McCabe, et al. 2008).

This thesis contributes to the understanding of computational model validation in three ways, by: conducting a two-part validation of a model of patient illness, generating a conceptual model so that explanations can be generalized from simulations, and developing an N=1 approach to validating meaningful variation over time in individual patients with chronic disease. The validation is a two-part study of the SimCare Patient Model. The first part is a conceptual validation that defines what aspects of a real-world problem are being modeled and why. How these aspects are represented in the model as sets of variables and functions is also defined. The conceptual validation provides

transparency as to the workings of the model, a basis for generalizing explanations related to model predictions or emergent behavior, and the relevant contexts for model utilization.

The second part is an operational validation that conducts two sets of simulation experiments to compare model predictions to observed values. Each set of experiments is used to characterize model accuracy in different contexts: The simulation of aggregated outcomes of cohorts of patients responding to treatment protocols in controlled trials and of meaningful variation in individual patients responding to treatments in a clinical care setting. The first set of experiments compares the simulated results of three published randomized clinical trials – each with a different focus on a main aspect of treatment of patients with type 2 diabetes – using three different cohort measures: nominal intermediate health outcomes, relative intermediate health outcomes and cardiovascular disease event rates. One trial has also been simulated by multiple, alternative type 2 diabetes models and provides a basis for comparison of these models with SimCare. The second set of experiments compares actual treatments and outcomes drawn from de-identified electronic health records in a clinical care database to a range of simulated responses from identical synthetic patients and treatments over the course of a year, one patient at a time (N=1).

The contributions of this thesis can be organized into three related parts, 1) a two-part validation study of a computational model of patient illness, 2) a conceptual model to be

used as the basis for generating explanations for model behavior, and 3) a novel form of operational validation using an N=1 experimental approach to measure meaningful variation in individual patients over time. The validation is presented to satisfy the interests of two overlapping research communities – those interested in the content of the model: the healthcare research community; those interested in computational modeling and validation techniques: the computer science community. The validation study is divided into a *conceptual* validation and an *operational* validation. The conceptual validation establishes the set of relevant theories identified in the natural system to be represented in the model. These theories enable the explanations of the model to be generalized and learned from, and they define the intent and contexts for relevant uses of the model.

The operational validation performs two types of simulation studies that characterize the outputs of the model under two different real-world contexts. The first set of experiments compares the simulation of populations of individuals under treatment protocols to the outcomes of three well-known clinical trials in the diabetes community. This distinguishes the model as being able to simulate controlled trials to the extent that a population of individuals can be generated and treatment protocols defined.

In the second set of simulation experiments, a series of N=1 trials are conducted using retrospective, outpatient clinical care data to demonstrate that SimCare accurately represents meaningful variation over time in individuals being treated for diabetes in a

clinical (i.e., less controlled) setting over time. In this setting, meaningful variation is defined as the non-random, clinically relevant variation in outcomes that can emerge over time given a specific course of treatment and an initial patient state. For example, if a physician were to treat two simulated patients via model software, and each patient had identical, observable initial states and received same treatments, the physician would not expect the two patients to exhibit identical responses to the treatments. This variation that exists in the real world of clinical care and causes an individual patient to exhibit meaningful differences in outcomes over time is an intentional part of the SimCare model and requires its own validation study. This distinguishes the model as being one of individual patients (rather than population-based) representing common, primary care encounters (rather than pre-screened patients under controlled protocols).

The results of the conceptual validation show that the SimCare model is clinically transparent and capable of generating explanations related to treatment outcomes. The operational validation shows that the SimCare model is able to capture meaningful and typical variation both in individual patients over time and across sets of patients in controlled cohorts.

Table of Contents

List of Tables

List of Figures

Introduction

> *"…one [student] once told me I was a model professor. I thought this was high praise until I realized that a model is a small imitation of the real thing."*
>
> *"All models are wrong."*
>
> *— John D. Sterman*

Perhaps Professor Sterman misspoke. Many students might remember their high school chemistry classes where displayed models of molecules were often quite large relative to their actual size in nature. It is not that a model is a "small imitation" but a simpler one. It is this simplicity that makes models feasible to work with and also what makes them "wrong." However, not all models are equally wrong or wrong for the same reasons. In the case of the molecule, the model was wrong because it didn't work at all like a molecule, but its tangible size and rigid structure made it a simple and effective teaching tool. A computational model of a molecule would be wrong for different reasons.

The tension between being simple and being wrong is at question in the process of model validation. Is the model in question simple enough to use, understand and communicate? Is it "right enough" in the ways it was purposefully designed? This tension can also be expressed between the explanation and prediction a model provides. The explanation a model provides is related to the reasons for and methods of its construction. When the model generates an interesting observation, does it have the ability to explain to the observer the what, why and how of the observation? The prediction a model provides is related to its ability to generate outputs, regardless of how it is accomplished. Even a

"black box" model could be useful in its ability to make predictions (though it would generate no explanation for them).

Computational modeling at its heart is the representation of something in sets of mathematical expressions. The selection of which mathematical expressions is an important step in constructing a computational model, as these expressions must both accurately represent the concepts identified in the model and satisfy practical, computational considerations. This can contrast with a purely mathematical model, which may not take into consideration implementation or experimental concerns but instead expresses a theory (e.g., $E = mc^2$). In the case of a computational model, mathematical expressions are selected for the purpose of being programmed into a computer, which acts as a catalyst for experimentation. Rather than manually performing thousands of computations over the course of several weeks to conduct one experiment, scientists have the luxury of performing thousands of experiments in an hour on personal laptops. Although mathematical modeling has long existed, this recent development in computational access to speed and precision – along with advances in artificial intelligence – has created a paradigm that is unique in history. The speed of innovation is exciting; the output of computational models can be seductive. But computational models are independent of their implementation on the computer technology with which we interface; it is the computer technology that is driven by the particular mathematical expressions of the model, and that set of expressions that is driven by the concepts being modeled.

A working software version of the SimCare Patient Model was originally implemented as a research tool to aid in the study of physician decision-making (Dutta, et al. 2005). The software was implemented as a rule-based, functional model used to generate successive patient health states. SimCare was designed to represent a patient with type 2 diabetes from the perspective of a primary care physician in an outpatient clinic and included patient state information about blood sugar, blood pressure, lipids panel and other relevant health indicators as they changed over time within the dynamics of chronic disease progression and related treatments. This computational tool enabled a variety of research involving the presentation of case-based, simulated patients to physicians and medical residents in order to research the effects of such interventions on physician decision-making.[*] Dynamic decision-making research and cost-effectiveness experiments using process control and control theory as generative models of physician decision-making have also been conducted by interfacing such models with the SimCare Patient Model.[†] These experiments were conducted to identify more- and less-successful decision-making strategies by observing the effects of resulting treatment patterns on simulated patient outcomes and then explaining such outcomes through the workings of both the SimCare Patient Model and the decision-making models. Additionally, the

[*] (Sperl-Hillen, O'Connor, et al., A Simulated Physician Learning Program Improves Glucose Control in Adults with Diabetes 2010) (O'Connor, Sperl-Hillen, et al., Simulated Physician Learning Intervention to Improve Safety and Quality of Diabetes Care: A Randomized Trial 2009) (O'Connor, Sperl-Hillen, et al., Customized Feedback to Patients and Providers Failed to Improve Safety or Quality of Diabetes Care: A Randomized Trial 2009) (Sperl-Hillen, O'Connor, et al., Personalized Physician Learning Intervention Improved Glocose Control in Adults with Diabetes 2009) (Sperl-Hillen, O'Connor, et al., A New Approach to CME? A Simulated Physician Learning Program Improves Glucose Control in Adults with Diabetes 2010)

[†] (Ramsey, Johnson, et al., Computational Models for Investigating Success and Failure in Treating Patients withType 2 Diabetes 2010) (Ramsey, Johnson, et al., Using Functional Data Analysis to Identify Physician Strategies which Lead to Better Type 2 Dabetes Patient Outcomes 2010) (Meyer, Adomavicius, et al., Towards Lower Macrovascular Risk in Diabetes Patients: A Simulation-Based Evaluation of Prioritization Strategies 2010) (Ramsey, Johnson and O'Connor, et al., Identifying Physician Decision Strategies for Treating Patients with Type 2 Diabetes 2010) (Gilmer, et al. 2010)

combination of both the decision-making models and the SimCare model created a widely distributed and detail-rich output of synthetic data that enabled data mining experiments – otherwise difficult or infeasible to conduct on available clinical data – to predict and classify a variety of patient outcomes and decision-making characteristics.[‡] With the growing body of research that uses the SimCare model prototype as an experimental test bed, a thorough validation had yet to be conducted (McCabe, et al. 2010).

In order to clarify the contributions of this thesis as well as credit previous research, an organization is presented as an overview of computational modeling. In Figure 1, a simplified version of the modeling process is presented (Sargent 2004). In such a view, the *problem entity* is the system – real or proposed – idea, situation, policy or phenomena to be modeled. The *conceptual model* is the logical representation of the problem entity developed for a particular study. The *computerized model* is the conceptual model implemented on a computer. These three components are related to each other by three phases (dotted lines). The conceptual model is developed through an *analysis and modeling phase*, the computerized model is developed through a *computer programming and implementation phase*, and inferences about the problem entity are obtained by conducting computer experiments on the computerized model in the *experimentation phase*. *Data validity* is defined as ensuring that the data used throughout these phases are adequate and correct.

[‡] (McCabe, et al. 2008) (Ramsey, Johnson and Adomavicius, et al. 2008) (Meyer, Adomavicius, et al., A Machine Learning Approach to Improving Process Control 2009)

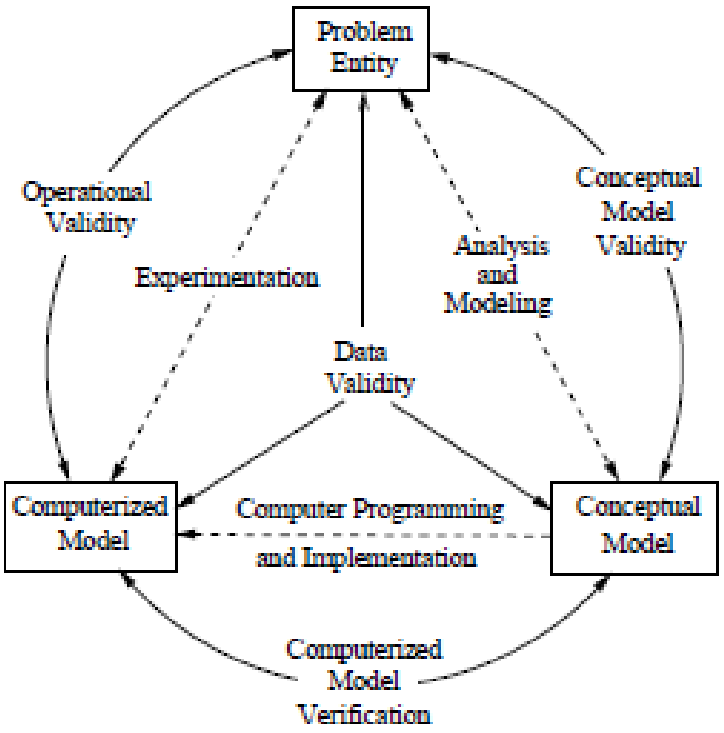


Figure 1. A simplified version of the modeling process

Model validation and verification are represented in this simplified figure by the lines encircling the diagram. *Conceptual model validation* is defined as determining that the theories and assumptions underling the conceptual model are correct and that the model representation of the problem entity is "reasonable" for the intended purpose of the model. *Computerized model verification* is defined as assuring that the computer programming and implementation of the conceptual model is correct. *Operational validation* is defined as determining that the model's output behavior has sufficient accuracy for the model's intended purpose over the domain of the model's intended applicability. Given this organization of the model development process, previous research contributed to the initial implementation of the computerized model whereas this

research contributes the conceptual model and two forms of model validation (operational and conceptual model validation).

The contributions of this thesis can be organized into three related parts, 1) a two-part validation study of a computational model of patient illness, 2) a conceptual model to be used as the basis for generating explanations for model behavior, and 3) a novel form of operational validation using an N=1 experimental approach to measure meaningful variation in individual patients over time. The validation study is divided into a *conceptual* validation and an *operational* validation. The conceptual validation establishes the set of relevant theories identified in the natural system to be represented in the model. These theories enable the explanations of the model to be generalized and learned from, and they define the intent and contexts for relevant uses of the model.

The operational validation performs two types of simulation studies that characterize the outputs of the model under two different real-world contexts. The first set of experiments compares the simulation of populations of individuals under treatment protocols to the outcomes of three well-known clinical trials in the diabetes community. This distinguishes the model as being able to simulate controlled trials to the extent that a population of individuals can be generated and treatment protocols defined.

In the second set of simulation experiments, a series of N=1 trials are conducted using retrospective, outpatient clinical care data to demonstrate that SimCare accurately represents meaningful variation over time in individuals being treated for diabetes in a

clinical (i.e., less controlled) setting over time. In this setting, meaningful variation is defined as the non-random, clinically relevant variation in outcomes that can emerge over time given a specific course of treatment and an initial patient state. For example, if a physician were to treat two simulated patients via model software, and each patient had identical, observable initial states and received same treatments, the physician would not expect the two patients to exhibit identical responses to the treatments. This variation that exists in the real world of clinical care and causes an individual patient to exhibit meaningful differences in outcomes over time is an intentional part of the SimCare model and requires its own validation study. This distinguishes the model as being one of individual patients (rather than population-based) representing common, primary care encounters (rather than pre-screened patients under controlled protocols).

The results of the conceptual validation show that the SimCare model is clinically transparent and capable of generating explanations related to treatment outcomes. The operational validation shows that the SimCare model is able to capture meaningful and typical variation both in individual patients over time and across sets of patients in controlled cohorts.

The three contributions presented in this thesis satisfy the interests of two overlapping research communities – those interested more in the content of the model: the healthcare research community; those interested more in computational modeling and validation techniques: the computer science community. The overall two-part validation of the model, the generation of the conceptual model and the N=1 validation technique

contribute to both research communities to different extents and for different reasons. To the healthcare research community, the validation of the model is itself a contribution directly related to a valuable, working model of a patient with type 2 diabetes. The conceptual model detailed within this thesis is an additional contribution which creates physician access to the representations therein as well as a resource to help explain model predictions. The N=1 validation should be viewed as a contribution because it uses heterogeneous clinical care data – generated outside of the control of research trials – and presents a formal validation technique that could be generalized to other chronic disease models.

The computer science research community may view the contributions slightly differently. Under the umbrella of computer science and computational modeling, model validation remains an area of ongoing research. The overall, two-part validation of the computational model is itself a contribution as an example of computational model validation in a multi-disciplinary setting. The particular focus on the conceptual model and conceptual validation may be viewed as an additional contribution to computer scientists, where often the focus on model validation is reduced to the accuracy of its predictions. In artificial intelligence some types of problems (e.g., in machine learning) can be thought of as a set of input data, a selected algorithm to learn a mapping and the resulting predictions of target variables. In this context, the computational model is a "black box" that is validated solely on its ability to predict something – the content of the model is strictly related to the computational aspects of the algorithm used and not explicitly related to the natural system being modeled. The conceptual model and its

validation are useful contributions for computer scientists, especially with artificial intelligence backgrounds, to consider additional types of computational models, wherein their design also includes the way computations are constructed for the purposes of providing explanation.

For the third contribution of the N=1 variation validation, much of the inspiration for its development came from some techniques in computer science used to understand the scalability of a set of computations as they scale with input size (N) asymptotically. One such expression is called "big O notation" and describes a way to estimate the complexity or behavior of a more complex function by using a simpler one. Other areas of computer science, in an attempt to estimate by analytical proof the computational complexity of a function, demonstrate how a similar but better known function can be used as an asymptotic boundary for the function under question. This approach may enable certain analytical actions to be performed to the simpler function that were difficult to perform to the original function, so that the proof may result in unequivocal statements about a given computational complexity. In the case of the N=1 validation in this thesis, the fundamental question was not to measure the accuracy of a specific prediction (e.g., a simulated patient) but to measure the accuracy of the possible variation for any given patient under a variety of heterogeneous and usually unknown assumptions. These unknown parameters are naturally bounded by reality – not theoretical asymptotes but distributions of parameters generated by biological characteristics of populations of individuals. This form of N=1 bounded validation is appropriate for the reasons associated with this model's intended usage and design, however computer scientists

could generalize this validation method to any computational model that is sampled over time (e.g., differential equations) but where values of modulating parameters in any specific simulation case are unknown but also bounded by the nature of the system (i.e., realistically non-infinite).

The thesis is organized into four chapters and a set of appendices. The first chapter details the conceptual validation of the SimCare model. The second chapter contains two sections of operational validation. The first section covers the three randomized clinical trial simulations used as the first part of the operational validation. The second section presents the N=1 clinical care operational validation of the model. The third chapter is a discussion and conclusion of the thesis. The appendices contain additional detail for various parts of the model including drug and dose formulary definitions and how a particular implementation of the SimCare model can be used to conduct simulation experiments.

Conceptual Validation of the SimCare Model

Simulation models have been used throughout history for such wide-ranging purposes as professional training (Issenberg, et al. 1999), military games (Pew and Mavor 1998), and scientific modeling (De Jong 2002). In any case, such simulation models are used to gain limited access to something more complex in the real world. How a model is "limited" depends upon the intent of the modeling exercise and the way the natural system is conceptualized.

A basic modeling relation, as described by Mikulecky (Mikulecky 2000), is between a *natural system* and a *formal system* (Figure 2). The natural system is something we wish to understand. The formal system is the way we encode that natural system. Mathematical models express concepts in the language of mathematics but are not necessarily intended to be implemented on a computer. Computational modeling is a specific kind of mathematical modeling that reduces a natural system to a formal system as a set of mathematical expressions. This formal system can then be implemented through a computer so that simulations can be run. The formal system for a computational model can be thought of as a combination of a conceptual model and an operational model. An appropriate validation study for each model – conceptual and operational – is conducted.

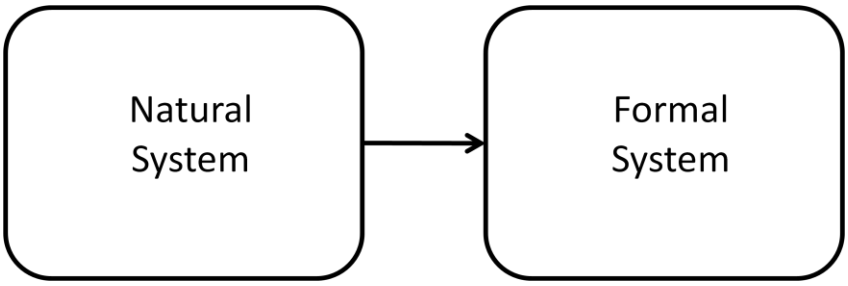


Figure 2. A simple representation of a basic modeling relationship. In the case of the SimCare model, the natural system is the individual patient receiving clinical care in an outpatient setting. The formal system is the SimCare model. The arrow represents the level of abstraction of the natural system used to encode the formal system.

For this thesis, conceptual validation is defined as determining that the model representation of the problem entity and the model's structure, logic, and mathematical and causal relationships are "reasonable" for the intended purpose of the model (Sargent 2004). This is a necessary part of model validation if the model is developed in order to provide *explanation*. Explanation and prediction are two aspects of the modeling problem that are often discussed in tandem; not every model does both. A model can be created in an effort to produce the most accurate predictions possible given a set of observations (e.g., Bayesian models) or can be constructed by a set of universal laws about the natural system (e.g., propositional logic). In order for a model to provide more meaningful explanation for predicted behavior, it must be more than a system of inputs and outputs or a set of deductive-nomological statements; it must be built in a way that actually represents the functioning in the natural system (Bechtel and Abrahamsen 2005). This gives the model a mechanistic property which must, itself, be validated. As is implied by the working parts of a conceptual machine, a simple checklist of parts is not enough to validate the workings of the model. Rather, a rich description about the

problem entity must be provided including any hierarchy of function or inter-dependent dynamics of the parts of the natural system being modeled.

David Marr has outlined an elegant framework for establishing the representation of natural systems through his work in the field of vision (Marr 1982). This framework has been used by different scientists in a variety of disciplines from social cognition (Mitchell 2006) to psychological cognitive science more broadly (Cosmides and Tooby 1995) (Gigerenzer and Hoffrage 1995) (Laville 2000). Marr argues for two types of relevant theories – Type I and Type II – to guide the formation of what defines the information processing problem (Marr, Artificial Intelligence -- A Personal View 1977).

A Type I theory first establishes a *computational theory* for the system by abstracting the problem formulation of "what" is being computed and "why". What specific problem relating to the natural system is being solved? What constraints enable this problem to be considered in this way and generalized? From the statement of the problem as an answer to what and why, the "how" is described in terms of the algorithms used to encode the formal system. Marr outlines a computational level, an algorithmic level and a physical level. Each level is independent of the other, though the ordering is in logical progression (i.e., the computational level informs the algorithmic level, etc.). Together, the what and the why form the computational level (or computational theory), and the how forms the algorithmic level. The physical level is the implementation level. In the case of the natural system, the implementation is the person with diabetes; in the formal system, the implementation is the computer code.

Part of the importance of constructing a Type I theory can be seen in contrast through the limitations of what Marr calls Type II theory. Type II theory can be thought of as the "how" of the system but without any statements as to what is being modeled and why. It consists of its own simplest description of the particular approach used to solve a problem. Type II theories are judged on their ability to make good predictions (e.g., some goodness of fit criterion) and lack the ability to provide a generalization or explanation of the problem. For example, a probabilistic model of physician decision-making could be created based on a statistical description of a database of treatments. This model might be constructed such that it accurately predicts physician moves over a group of patients, but nothing could be learned from this model concerning the actual decision making of physicians – the model would lack any principled description of the process of decision making.

In the case of the SimCare Patient Model the creation of a Type I theory provides a foundation for generalization and understanding patient's responses to treatments for type 2 diabetes. Marr's framework for a Type I theory is followed throughout the conceptual validation of the SimCare model and consists of two parts: the computational level and the algorithmic level.

The first part is the level of abstraction (the "what" and the "why") used to frame the system being modeled in terms of the problem that is being solved. The problem being solved by the SimCare model is the succession of responses over time exhibited by

individual patients undergoing treatments for type 2 diabetes while under clinical supervision of a primary care physician. This problem is organized by two constructs: the representation of diabetes in an individual and the representation of the clinical encounter over time. The second part of the conceptual validation describes the represented processes that generate patient states and responses to treatments. The algorithmic level is described in terms of patient state variables, actions that can be taken in the context of providing clinical patient care, and the pathways that express change in these variables over time.

An important part of each level in Marr's analysis is identifying the *constraints* of the problem. For example, as Mitchell pointed out in the field of social cognition, computational theory includes those constraints of the model that are both *useful* and *useable* (Mitchell 2006). Useful means that the information in the SimCare model is a reliable representation of the actual state of the patient (i.e., that could be acted on by a physician problem solver). Useable means that the information in the model is, in practice, available in the real world. For example, blood glucose (represented in SMBG or Hemoglobin A1c) is a measure of the patient's state that enables a physician problem solver to take action (useful) and is available to a physician in the clinic upon request (useable). Other constraints, such as patient adherence to treatment, are useful in this problem but not necessarily useable – not directly available in a clinical setting. Such constraints allow the model to be evaluated in relation to the natural system in question as well as generalized to other systems. The constraints detailed in the SimCare model have been developed from medical literature and from input from physicians dealing with the

immediate context of patient care. As understanding of patient care evolves, additional constraints are added and existing ones modified. Throughout this thesis, a "constraint" describes a property of the natural system being modeled, a "feature" describes its formal representation in the model, and a "variable" describes the representation at the algorithm level.

The algorithm level theory is generally constrained by the characteristics of the implementation level. As in Marr's ubiquitous example, if one is designing an algorithm for adding the prices of retail goods at a store for purchase (e.g., as in a cash register), then the algorithms used to do this would be different if the hardware available to implement the algorithms was not a computer but an abacus. For the SimCare model, the computational and algorithmic theory levels are related to their natural implementation as a real patient with type 2 diabetes. This validation generalizes the implementation level of the formal system – which was developed previously as an implementation of computer code – to the levels of algorithmic and computational theory.

Part I: The Conceptual Level – "What" and "Why"

The SimCare Patient Model represents patient states and changes in patient states from treatments that are provided in the context of outpatient clinical care. The computational theory is defined by two constructs: 1) the way type 2 diabetes is manifested in an individual patient, and 2) the way the individual patient is represented for purposes of treatment in the outpatient clinical encounter.

Each construct establishes sets of constraints that define the computational theory of the patient model. The first construct distinguishes the SimCare model from other forms of type 2 diabetes models (e.g., epidemiological, etc.) and from models of patients with type 1 diabetes. Such models have been designed to solve other problems. The second construct distinguishes the way the SimCare model conceptualizes patients undergoing treatments for type 2 diabetes in sequences of patient-physician encounters in a clinical setting. Other models (of type 2 diabetes) have been designed from the perspective of population treatments or pathophysiological representations of a patient and are less suited to address problems of patients responding to treatments by clinical care physicians over time. The following two sections detail the constraints in each construct of the SimCare model.

Construct 1: Type 2 Diabetes

Type 2 diabetes, as a disease, is the first construct expressed in the SimCare model. The constraints that reflect type 2 diabetes in a patient with this condition can be organized in terms of the *observable* and *unobservable* patient states plus *responses to physician actions* that cause changes in these states over time (Figure 3). Observable patient states consist of features that are accessed by the physician problem solver in the clinical setting. Unobservable patient states consist of features of the patient that are not directly accessed by the physician. The physician actions change values of features and generate relevant patient responses.

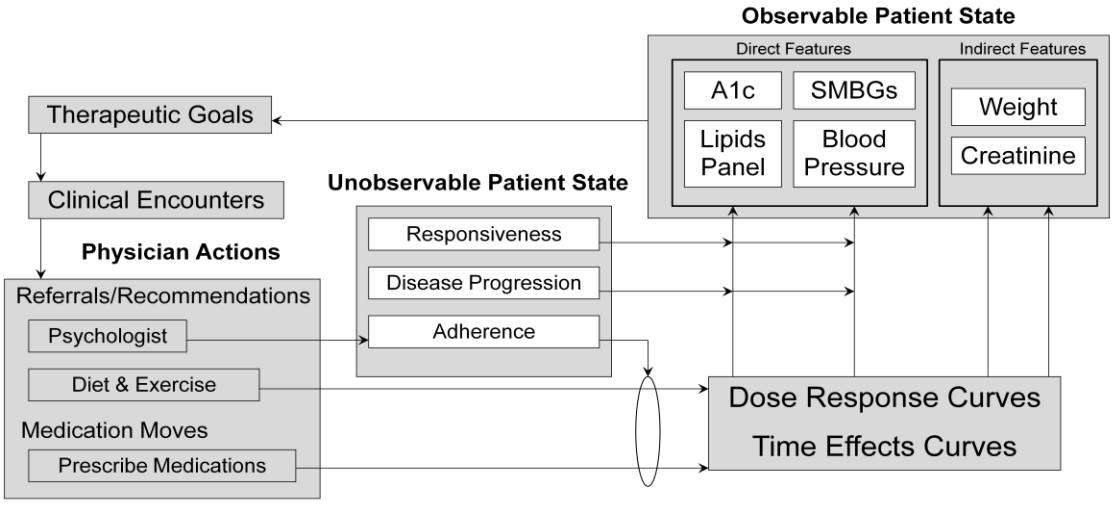


Figure 3. An overview of the SimCare Patient Model and its clinical context. The constraints of the SimCare Patient Model are expressed as features of patient states (white boxes). Arrows represent relationships among features and other parts comprising the construct of type II diabetes. The direction of an arrow indicates the direction of influence between parts. The patient state can be divided into observable and unobservable features. The observable patient state can be queried for values at any point in time and is affected directly and indirectly by physician actions. The unobservable patient state modulates the affects that physician actions have on the observable state and can also be affected by some physician actions.

Type 2 diabetes is directly a dysfunction of blood sugar and insulin balance in the body (Dutta, et al. 2005). Unlike type 1 diabetes where insulin is not produced by the body and, therefore, must be injected to regulate blood sugar levels, type 2 diabetes reflects a range of imbalance between insulin resistance and high blood sugar. A variety of outpatient clinical options exists for the management of type 2 diabetes. Typical approaches to outpatient clinical treatments of type 2 diabetes include modification of diet and exercise, the use of single or multiple oral hyperglycemic drugs (e.g., Metformin) and, in some cases, the use of insulin. Insulin as a treatment is fraught with difficulties. It can induce weight gain in patients, increase the susceptibility of hypoglycemic events (e.g., low energy, fainting) and can decrease adherence levels due to patient fear of needles, thus insulin use in patients with type 2 diabetes is often considered a treatment of last resort.

In addition to being associated with obesity, type 2 diabetes is also often associated with high blood pressure and high lipids levels. These multiple conditions each contribute to higher levels of potentially serious comorbidities – such as stroke, heart attack, kidney failure and blindness – and must be represented in the model in order to provide information that is used in determining relevant treatments for a patient with this condition.

In prior work, an implementation of the SimCare model was constructed using medical literature and a team of consulting physicians. Thirteen key health indicators of a type 2 diabetic patient were identified and represented: Hemoglobin A1c, self-monitored blood

glucose (SMBG), systolic blood pressure (SBP), diastolic blood pressure, low density lipoproteins (LDL), high density lipoproteins (HDL), triglycerides, serum creatinine, weight, body mass index (BMI), height, adherence to treatment, and depression (Dutta, et al. 2005) (Stamler, et al. 1993) (Turner, et al. 1998) (Standl, et al. 1996).

These features represent a set of intermediate health outcomes that contribute to microvascular and macrovascular disease[§] (e.g., renal function and blood pressure, respectively) and may contraindicate specific treatments (e.g., high serum creatinine levels and Metformin). At any point in time (daily) the patient state is defined by values for each of these features.

Observable Patient State

Observable states of the patient model are divided into *direct* and *indirect* sets of features. Direct features are those intrinsically affected by the disease and are targets of medications used to treat the disease. Indirect patient features are also affected by medications but as side-effects of the treatments of direct features. The indirect features are no less important, however, as they can greatly impact the overall health of a patient or contraindicate specific medications. For example, some medications (such as lisinopril) can cause creatinine levels to rise in patients, and other medications (such as metformin) can be contraindicated by high levels of creatinine. Creatinine levels,

[§]Avoiding or delaying the onset of microvascular and macrovascular events in patients is a primary concern that guides treatment options and evidence-based medicine. Micro- and macrovascular events are not generated by the SimCare model directly but can be simulated by using an additional risk model.

themselves, are not directly targeted for treatment of patients with type 2 diabetes, however failure to recognize creatinine levels is a serious error of commission (acting against a contraindication) in patient care.

Direct Features

Direct features of the observable patient state in the SimCare model are *blood glucose*, *blood pressure* and *lipids panel*. Blood glucose is represented in two ways: Hemoglobin A1c and Self-monitored blood glucose (SMBG). SMBG panels can show values for the three days previous to the current clinical encounter at as many as five times per day: 3am, pre-breakfast, pre-lunch, pre-dinner, and bedtime. HgbA1c (hemoglobin A1c) is a measure that represents an average of the previous 90 days of SMBG values.

SMBGs are the actual measurement of blood sugar levels in a patient at a given point in time (taken by finger prick) and are a significant component of gauging the patient's response to treatments (especially insulin treatments). Blood sugar in a patient is naturally volatile and can change rapidly from eating meals or the injection of insulin. Hemoglobin A1c is an indirect and longer-term measure of average blood sugar levels in a patient over time. A1c levels are taken by blood draw and measure the percent of glycated red blood cells in a patient. Red blood cells become glycated over time because of excess blood sugars. These red blood cells have a natural life cycle of approximately 90 days (Saudek, Derr and Kalyani 2006). Thus, the A1c level for a patient at a given time corresponds with the average blood sugar for that patient over the course of the past

90 days. The SimCare model represents these time effects differentials for the relevant SMBG and A1c values in a given patient's response to treatments.

Blood pressure is represented in the patient state in both systolic (SBP) and diastolic (DBP) form. A patient's lipids panel is made up of three independent components: LDL, HDL, and Triglycerides. Elevated LDL levels are a strong independent predictor of increased cardiovascular risk, and are often found in patients with type 2 diabetes. HDL and triglycerides are separately represented. Elevated HDL levels are an independent predictor of decreased cardiovascular risk. Elevated triglycerides levels also increase cardiovascular risk, but the independence of this effect is a matter of ongoing debate in the medical community. In the case of blood pressure, systolic and diastolic levels conjointly predict cardiovascular risk, whereas the components of the lipids panel act independently from each other.

Indirect Features

The indirect features of observable patient states are *body-mass index* (BMI), *weight*, *height* and *serum creatinine*. BMI is a computed feature from height and weight. Height is a constant for a given patient, therefore changes in BMI are a function of changes in weight. Weight is an important feature because it not only impacts the patient's health state but is used in developing treatment policies (e.g., for insulin). Most types of insulin not only cause a patient to gain weight, but weight itself is a main component of adjusting insulin doses. Otherwise, weight gain does not directly affect other features of the patient

state. Serum creatinine is represented because it signals an overly taxed kidney and can be an important contra-indication for certain medications.

Unobservable Patient State

Some features of the patient state are not directly measurable at the clinical encounter but are represented because they also affect changes in patient states. These features introduce nonlinearities (patient variation) both within a given patient over time and across patients. Unobservable patient state features in the SimCare model are *responsiveness to drugs*, *disease progression* and *adherence*.

Physical responsiveness to drugs represents variability in responses to same treatments among otherwise similar patients. It is represented separately for each drug and each observable feature based on published reports of randomized clinical trial data and drug company data[**]. Another biological feature that differentiates otherwise identical patients is disease progression. Disease progression models the way an untreated patient state continues to change over time in the absence of interventions. Disease progression is modeled in relationship to each observable patient feature independently based on published reports of clinical trials (Levy, et al. 2004) (Cowie and Harris 1995). Adherence to treatment models a psychosocial aspect of the type 2 diabetic patient and is affected by a variety of factors, including the ability to afford medications, belief in the seriousness of the disease, and the willingness to comply with treatment protocols (Rubin

[**] (Campbell 1988) (DeFronzo 1999) (Garber, et al. 1997) (Aronoff, et al. 2000) (Lebovitz HE, et al. 2001) (Esposito, et al. 2008)

2005) (Ciechanowski, et al. 2001) (Dunbar-Jacob and Mortimer-Stephens 2001). This feature is applied to all forms of treatment of the patient state from diet and exercise to patient medication prescription fills.

Responses to Treatments

An important constraint to represent in the patient model is access to the disease features of the patient state that are typically available to physician problem-solvers. The combination of patient state features (both observable and unobservable) and physician access to patient state information are the constraints that represent how the patient state responds to treatments at a point in time. The constraints that define the responses to physician actions are the treatment options of *referrals/recommendations* and *medication moves*.

Referrals and recommendations generate different types of response in the patient state. Referral to a psychologist can increase the adherence level of the patient (via treating depression) and can therefore increase the effects of other treatments acting upon the patient. Improving diet and exercise, usually through referral to a nurse educator, can also be recommended by a physician and can impact features of the patient state such as blood sugar, blood pressure, lipids, weight as well as unobservable features like adherence. Changes in these features of the patient state are represented through increased exercise levels and dietary changes such as decreasing fat, calorie and sodium intake while increasing fiber intake.

Medication moves impact clinical features in the patient state. Hyperglycemia (elevated blood sugar) is treated by families of oral medications (metformin, sulfonylureas, TZDs, etc.), GLP-1 receptor agonists, and different types of insulin. Lipids (HDL, LDL, and Triglycerides) are treated by families of statins and fibrates. Blood pressure medications of various classes (e.g., diuretics, angiotensin converting enzyme inhibitors, beta blockers, calcium channel blockers, angiotensin receptor blockers, etc) are used to lower systolic and diastolic blood pressure. Oral medications can be administered *q.d.*, *b.i.d.*, *t.i.d.*, and *q.i.d.* (i.e. one, two, three or four times daily), depending upon the pharmacokinetics and duration of effect of various medications. Responses to medications are represented independently so that responses can be generated when new drugs are introduced to the system, when multiple drugs are given or when drugs are removed from the system.

Side-effects from medication moves appear in indirect features of the patient state. For example, some blood pressure treatments, in addition to lowering systolic blood pressure, tend to raise serum creatinine levels. This is a relevant aspect of clinical care because high levels of serum creatinine indicate an over-taxed kidney and contraindicate the use of a common and effective blood sugar treatment medication called Metformin.

The magnitude of patient state responses to treatments depends upon a combination of quantity and timing. Quantity of response is modeled through dose response curves. Dose response curves provide accurate responses to a variety of physician actions taken

(e.g., different doses of medication). These curves were taken from published randomized trials[††] and can be adjusted based on new information from ongoing trials.

The time elapsed until the next patient state is generated also impacts the magnitude of response in the patient state. Some medications take weeks or months to fully affect the patient state. These timing effects are modeled independently for each physician action as a series of time effects curves. This combination of dose and time effects curves allows an accurate patient state to be generated in response to any combination of physician actions over any sequence and amount of time. This timing constraint is of particular importance as most treatments have a delayed effect on patient states (Dutta, et al. 2005).

Construct 2: The Clinical Encounter

Type 2 diabetes is a multi-faceted, chronic disease often involving treating blood sugar, blood pressure and lipids levels. The features of the SimCare model reflect each of the patient-state constraints of the disease and its responses in a patient to treatments. Unlike acute care, chronic disease care is dependent upon successive treatments over time in order to achieve and then maintain a patient state of health. Part of the research problem posed by the SimCare model includes the representation of the context of the clinical environment – where the type 2 diabetic patient is treated. The disease of type 2 diabetes is represented by feature sets of observable patient state, unobservable patient state, and

[††] (Campbell 1988) (DeFronzo 1999) (Garber, et al. 1997) (Aronoff, et al. 2000) (Lebovitz HE, et al. 2001) (Esposito, et al. 2008)

responses to physician actions; the individual patient is conceptualized through the construct of the *clinical encounter*.

In order to represent patient responses to treatments over time, the construct of the clinical encounter was defined. This construct enables the intended use of the model to select treatment protocols and assess their impact on the patient state within the context of a clinical care environment. The three constraints of the clinical encounter are 1) providing *patient state information* for the physician problem solver, 2) *responding to physician actions* from previous encounters, and 3) *responding to the timing* of sequences of encounters.

Patient State Information

The problem of providing information about the patient state reflects the clinical reality that the patient state will remain unknown unless tests are performed (i.e., unless the patient state is queried). If tests are performed, then information about the patient state must reflect the current value of each health characteristic in the same manner as a lab test in a healthcare clinic (i.e., not an average value reported over the course of a year but a current, temporal value). This includes current A1c, SMBG (self-monitored blood glucose), systolic blood pressure and lipids panel (LDL, HDL, triglycerides). Test ordering has a different effect upon the information delivered depending upon the test involved. For example, A1c and lipids tests must be ordered in advance (e.g., at the previous encounter), whereas blood pressure tests are available in the clinic at the time of the encounter. The patient model can provide SMBGs at as many as five times per day

(pre-breakfast, pre-lunch, pre-dinner, bedtime, 3am) for the three days prior to the current, represented clinical encounter.

Response to Physician Actions

Treatments can be the initiation, titration (increasing a current dose) or reverse-titration of medications or a referral to a specialist or nurse educator for the purpose of improving diet and exercise. Each treatment has an effect on the patient state; this effect may or may not be delayed by the time course of the treatments. For example, the full blood glucose lowering effect of the oral medication Metformin does not manifest in the patient right away; it will take up to 14 days until the full effect manifests in the patient's blood glucose (SMBG) levels and up to 90 days to show up fully in the patient's A1c level (DeFronzo 1999). Furthermore, the time to full effect of each available medication and dose is not the same. Time to full effect is typically slower with medical nutrition therapy than it is with medications. Various medications may also have widely different time to full effect – insulin lowers blood glucose within minutes to hours, depending on the type of insulin, whereas certain oral medications (e.g., TZDs) may take many weeks to exert their full biological impact and even longer, as discussed below, to fully impact A1c. The magnitude of various therapies on blood glucose and other features (e.g., BP, LDL) also differs.

Medications are given in specific doses and may be combined at a given encounter. The patient state must respond to each medication given at each dose, including medications

that are removed from the patient's system. The patient state must also respond to referrals made (e.g., psychologist, medical nutrition therapist) and manifest relevant responses in subsequent encounters. The combination of these forms of treatment is a central component of the patient model. For these reasons, the model has been constructed to generate patient states over non-continuous but discrete points in time. In other words, patient states need to be generated specifically to represent the patient at clinical encounters but do not need to be continuously maintained during the time between encounters. This allows a functional approach (e.g., sampling specific values from dose and time effects curves) to representing patient state responses in the SimCare model.

Response to Timing of Encounters

Scheduling determines how soon the patient will be seen again and can occur over any number of days (i.e., the smallest unit of time represented by the patient model is one day). There is no upper limit on the number of days specified between clinical encounters. The duration between consecutive clinical encounters is important for two reasons: 1) it determines the magnitude of treatment effects on the patient state at a point in time (e.g., the next encounter), and 2) it affects the treatment decisions of physicians. For example, if two consecutive encounters are close together, the full impact of any medication administered during the first encounter may not show up in the patient state by the time of the second encounter. This uncertainty in effect may delay the decision to make an additional medication move (for fear of inducing hypoglycemia). Likewise, if

encounters are far apart (e.g., one year), then the effect of a single medication move on the next patient state may be dominated by naturally occurring disease progression.

The clinical encounter construct establishes a framework for the patient model (Figure 4) and creates a clear contrast between the SimCare model and other models of type 2 diabetes. The model reflects the constraint that from the point of view of the patient or clinician, the patient state only becomes apparent during clinical encounters. These encounters represent opportunities to get information, adjust/start medication, make referrals and schedule the next encounter. The amount of time that passes in between consecutive encounters is reflected in the changes of the subsequent patient state but is not represented continuously.

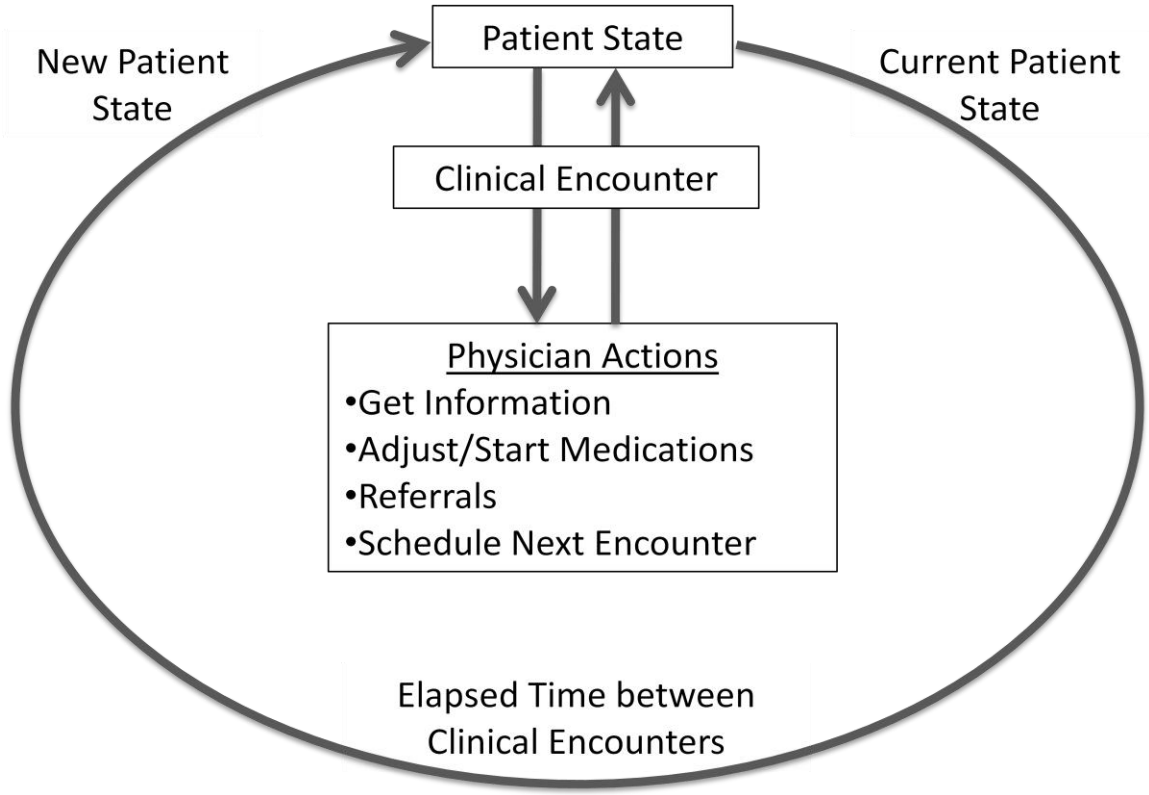


Figure 4. An intent of the SimCare Patient Model is to enable study of physician decision-making and is reflected in the construct of the clinical encounter. The construct is defined by the three constraints of providing: 1) patient state information for the physician problem solver, 2) patient state responses to physician actions, and 3) responses to the timing of sequences of encounters. These constraints can manifest differently according to the patient state feature in question, for example, slow-moving A1c or fast-moving SMBG levels.

As mentioned previously, other clinical models may reflect the population effects of treatments (e.g., epidemiological models) or the underlying pathology of disease (e.g., pathophysiological models). In the case of epidemiological models[‡‡], no attention is given to modeling specific individuals over time. Such models are designed to study the effects of general treatments on populations of patients, usually in annual blocks of time.

[‡‡] (Bagust, et al. 2001) (Clarke, et al. 2004) (Stevens, et al. 2001) (Muller, Maxion-Bergemann and Bolinder, et al. 2004) (Muller, Maxion-Bergemann and Gultyaev, et al., Development and validation of the Economic Assessment of Glycemic Control and long-term effects of diabetes (EAGLE) model 2006) (Muller, Maxion-Bergemann, et al., EAGLE diabetes model: basic features and internal validation of simulating long-term diabetic outcomes and related costs (Abstract) 2004) (Brandle and Herman 2004) (Palmer, Roze, et al., The CORE Diabetes Model: Projecting long-term clinical outcomes, costs and cost-effectiveness of interventions in diabetes mellitus (types 1 and 2) to support clinical and reimbursement decision-making 2004) (Earnshaw, et al. 2002) (Hoerger, et al. 2004) (Herman, et al. 2005)

In the case of the pathophysiological models[§§], the underlying biology such as organs and organ systems is modeled in detail to simulate patients and the effects of treatments. These models encode a more complicated internal system of concepts at the level of individual patient biology (e.g., more granular timescale) in order to study the underlying causes of disease and perhaps lack some of the important outpatient clinical representation. In the two comparison examples, each reflects type 2 diabetes in some way, yet the SimCare Patient Model represents the outpatient clinical encounter in an explicit way that is omitted by the other models.

Evidence-based Medicine

The SimCare model represents the outpatient clinical encounter to reflect an important constraint in evidence-based medicine: How do various, real-world treatment protocols (often translated from general, clinical guidelines) affect individual patients over time? Evidence based medicine is the use of current best evidence in making decisions about the care of individual patients (Sackett, et al. 1996). Current evidence and guidelines suggest individualizing patients' A1c goals in the range of < 6.5% to < 8%, depending upon comorbidity, likelihood of hypoglycemia, cognitive status, and other factors (Association 2011). For example, a goal for many patients' A1c levels is less than 7.0% (T. A. Group 2008). When patients with diabetes are first diagnosed, they are often discovered to have A1c levels higher than this goal level (Hu, et al. 2002). Even after

[§§] (Eddy and Schlessinger, Archimedes: A trial-validated model of diabetes 2003) (Eddy and Schlessinger, Validation of the Archimedes Diabetes Model 2003)

several years of treatment, many patients with an elevated initial A1c may not have achieved their goal level, despite the medical care provided (Eliasson, et al. 2005).

Several aspects of type 2 diabetes, apart from the nature of the disease itself, may affect a patient's ability to reach evidence-based goals. One is the lack of opportunities for a clinician to start or adjust medications. This is not related to the pharmacological effects of the treatments themselves but to the sparse scheduling of encounters. Another aspect is the reluctance on the part of the primary care provider to increase the number of medications or the doses of current medications if encounters are scheduled closely together. As described before, the time effects of disease progression and treatments are represented such that the patient state may not be accurately represented to the clinician due to delays in feedback or a lack of tests ordered. Finally, as described in the construct of type 2 diabetes, diabetic patients are often affected by additional diseases of blood pressure and lipids panel (Stamler, et al. 1993) (Turner, et al. 1998) (Standl, et al. 1996). The effects of clinical prioritization and personalization of the treatments of such "complex" patients remains a growing interest in the medical community (Gaede and Pedersen, Intensive Integrated Therapy of Type 2 Dabetes: Implications for Long-Term Prognosis 2004).

Part II: The Algorithmic Level – "How"

The SimCare model is of an individual patient with type 2 diabetes. The model was developed to represent a patient's response to treatments and disease factors over time. Additionally, the model captures the effects of not treating patients. After a patient state is initiated, a subsequent update of the patient state reflects changes in intermediate health outcomes due to disease progression and random variation over time.

The patient is represented by the values of the set of features in the patient state. These features comprise the status of the patient model at a given point in time (e.g., at a clinical encounter) and are generated based on prior patient states and the defined pathways that impact each feature (Figure 5). The represented constraints in the SimCare model have been detailed in the previous section (Part I) through the constructs of *type 2 diabetes* and the *clinical encounter*. The following section defines the variables used to encode the features and the variable pathways that generate successive values of a patient state.

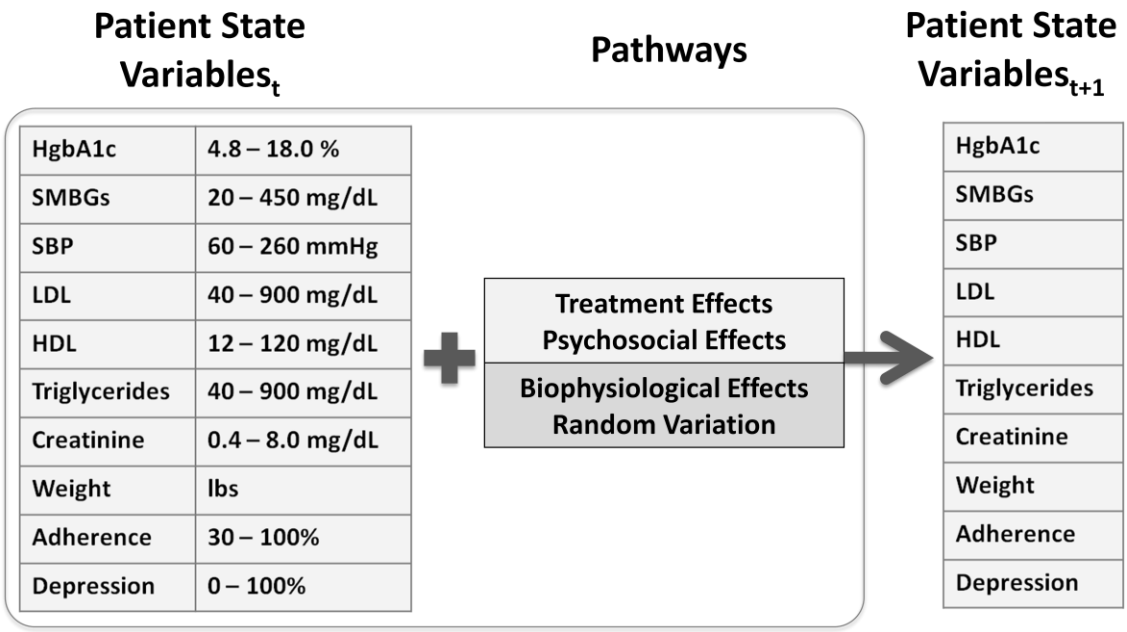


Figure 5. A new patient state is generated by applying the effects of pathways to each variable from the previous patient state. Aspects of pathways shown in the top box are directly affected by physician actions; aspects in the lower box are unaffected by actions but affect patient state variables. Ranges for the values of each variable are shown.

The complexity of the SimCare model can be described by the way the computations scale with inputs. The algorithms of SimCare scale with respect to the number of encounters simulated not the time over which the encounters span. For example, a simulation of 20 clinical encounters over the span of 20 years involves the same computational complexity as 20 encounters simulated over the span of 1 year. This reduces the complexity of the implementation and enables the use of combinations of functional algorithms. By contrast a model such as Archimedes – a disease pathology model – has a complexity in terms of the amount of continuous time the disease is simulated in a patient, in addition to the number of encounters. The representation of the inter-relationships of various organs and organ systems requires that encounters are linked continuously through time. SimCare, by querying functional combinations of dose

and time response curves, can generate a next state regardless of the elapsed time between states.

Patient State Variables

The constraints incorporated into the SimCare model are a set of intermediate health outcomes that contribute to subsequent microvascular and macrovascular disease[***] (e.g., renal function and blood pressure, respectively). At any point in time the model is defined by the set of patient state variables shown in Table 1.

Table 1. Patient State: Constraints and Variables

		Model Boundaries		
Constraints	**Variables**	**Min**	**Max**	**Unit**
Blood Glucose	A1c	4.8	18.0	%
	SMBG	20	450	mg/dL
Blood Pressure	Systolic Blood Pressure	60	260	mmHg
	Diastolic	40	160	mmHg
Lipids	LDL	40	900	mg/dL
	HDL	12	120	mg/dL
	Triglycerides	40	900	mg/dL
Renal Function	Serum creatinine	0.4	8.0	mg/dL
Body Mass	Weight	-	-	Kg
	Body mass index (BMI)	-	-	kg/m^2
	Height	-	-	*cm*
Adherence	Adherence	30	100	%
Psychosocial	Depression	0	100	%

[***] Microvascular and macrovascular patient health variables are used to guide treatment options and to capture intermediate health outcomes. Micro- and macrovascular events are not generated by the SimCare model directly.

The state variable of height is represented in *italics* and is fixed for a given patient throughout the duration of treatment; it is not affected by model pathways. All other state variables are affected by model pathways, which consist of physician moves (treatment and referrals), natural disease progression, adherence to treatments, responsiveness to specific drugs, weight fluctuation, and random variation.

Patient State Pathways

The variables in the patient state are computed based on the effects of pathways variables. Not all state variables are affected by each pathway variable. Each patient state pathway can be generally stated as:

Patient state = f(Prior Patient States, Treatment Effects, Psychosocial Effects, Biophysiological Effects, Random Variation Effects)

Prior patient states. Prior patient states hold the complete medical history of the patient up until the current point in time, when a new state is computed. Any changes in patient state variables that were not included in the current state computation due to modeled timing delays are included in the computation of subsequent states.

Treatment effects. Treatment effects model the interactions between the physician and the patient and consist of information-seeking, treatment and scheduling. Foremost, the model allows the use of medications and referrals to treat complex patients over time. Interactions of the patient model with changing doses of medications are affected by

other variables describing the patient.[†††] Each medication's effect is computed from specific dose and time-effects curves (described in detail in the Pathway Variables section).

Psychosocial effects. Several variables capture relevant effects in patient lifestyle and outlook outside of the clinic that affect the current patient state. These include adherence to treatment, depression, and changes in a patient's physical activity or diet.

Biophysiological effects. Without treatment, type 2 diabetes patients suffer from rising blood glucose, blood pressure and low-density lipoproteins (cholesterol); all factors which increase the risk of cardiovascular events (Turner, et al. 1998). This natural disease progression is modeled as a daily increase of the affected patient state variables. Additionally, individual responses to pharmacological treatments are adjusted to account for the biophysical individuality of the patient.

Random variation. Random variation in the SimCare model captures two aspects of patient states in the clinic. The first is the normal biophysiological fluctuations that occur over time (Boland, et al. 2001), and the second is the analytical error associated with measuring the patient state (e.g., A1c test measurement error) (Tran, et al. 2003). Both sources of variation in patient state are included in this variable and are modeled from published studies. This form of variation within a patient is applied to the patient state as it is generated for presentation at the current clinical encounter.

[†††] Patient Adherence levels reduce the effects of treatment moves by physicians. Individual physiological responsiveness to specific drugs also reduces or enhances the resulting effect for a given patient instance of the model.

Pathway Definitions

The following section defines the pathway for each variable in the patient state. The method of using dose and time response curves to compute the listed effects on each pathway variable is defined in the next section, Pathway Variables. (A complete formulary description of the dose response curves for each medication is found in Appendix A.)

Blood Glucose: Self-Monitored Blood Glucose (SMBG) values are measured by applying a drop of blood (obtained by the patient using a lancet) to a small portable sensing device called a glucometer. This important measure gives the patient a momentary reading of the amount (in mg/dL) of glucose in their blood and is often recorded several times a day to guide administration of glucose-lowering medications, especially insulin. A difficulty with interpreting SMBG values is that they can change within minutes due to meals or insulin shots.

Glycated hemoglobin (A1c) is a blood test that assays the percentage of hemoglobin molecules within red blood cells that are irreversibly bound to glucose. This percentage is reflective of the concentration of glucose in the blood over time – the more glucose in the blood, the more hemoglobin molecules are bound to glucose, and the higher the A1c value. Because red blood cells have about a 90 day life cycle, A1c reflects the moving average of blood glucose concentration over a 10-12 week period antecedent to the time of the A1c test.

If a patient has had no new medication introduced into their system within the past 90 days, then their A1c value will correspond to their average SMBG values for the past 3 months. If new medication is introduced into the patient's system, then it will affect SMBG values quickly and will have a delay in showing up in the patient's A1c. Modeling this dynamic is a necessary component to maintaining a clinically-plausible patient and is a crucial feature of the SimCare model.

Because SimCare is not a pathophysiological model, the biological pathway between blood glucose concentration and A1c is not explicitly modeled. The model does encode the information in these variables by modeling them using the same – although scaled[‡‡‡] – dose response curves for hyperglycemia treatments. They are maintained separately by using different time response curves such that A1c values accurately trail SMBG values through time. A1c and SMBG values are affected by the following variables:

$$A1c_{current} = f(A1c_{previous},$$

Adherence*[*OralMedicationEffects(*Responsiveness||*{Metformin, Glipizide, Glyburide, Glimeperide, Pioglitazone, Rosiglitazone}, InsulinEffects(*Responsiveness||*{Glargine, Lispro}), ExerciseEffects, DietEffects*], *DiseaseProgression, RandomVariation)*

[‡‡‡] The scaling factor between A1c and blood glucose is 44.4; this has been derived from a regression model of the HP population (regression model: A1c = (SMBG-87)/44.4 + 5.8) (Rohlfing, et al. 2002).

$SMBG_{current} = f(SMBG_{previous}$,

Adherence*[*OralMedicationEffects(*Responsiveness||*{Metformin, Glipizide, Glyburide, Glimeperide, Pioglitazone, Rosiglitazone}), InsulinEffects(*Responsiveness||*{Glargine, Lispro}), ExerciseEffects, DietEffects*], *DiseaseProgression, RandomVariation)*

Blood Pressure: Systolic and Diastolic blood pressures are affected by the following variables:

$SBP_{current} = f(SBP_{previous}$,

Adherence*[*OralMedicationEffects({Amlodipine, Atenolol, Lisinopril, Valsartan, Hydrochlorothiazide, Cardizem}), ExerciseEffects, DietEffects]*, *DiseaseProgression, RandomVariation)*

$DBP_{current} = f(SBP_{current})$

Lipids: LDL, HDL and Triglycerides are affected by the following variables:

$LDL_{current} = f(LDL_{previous}$,

Adherence*[*OralMedicationEffects({Simvastatin, Fenofibrate, Atorvastatin}), ExerciseEffects, DietEffects*], *DiseaseProgression, RandomVariation)*

$HDL_{current} = f(HDL_{previous},$

Adherence*[*OralMedicationEffects({Simvastatin, Fenofibrate, Atorvastatin, Gemfibrazol}), ExerciseEffects, DietEffects*], *DiseaseProgression, RandomVariation)*

$Triglycerides_{current} = f(Triglycerides_{previous,}$

Adherence*[*OralMedicationEffects({Simvastatin, Fenofibrate, Atorvastatin, Gemfibrazol}), ExerciseEffects, DietEffects*], *DiseaseProgression, RandomVariation)*

Renal Function: Serum creatinine is affected by the following variables:

$Creatinine_{current} = f(Creatinine_{previous},$

Adherence*[*OralMedicationEffects({Lisinopril, Valsartan})*], *RandomVariation)*

Body Mass: Weight and body mass index are affected by the following variables:

$Weight_{current} = f(Weight_{previous},$

Adherence*[*ExerciseEffects, DietEffects*] , *RandomVariation)*

$$BMI_{current} = f(Weight_{current}/Height_{current})$$

Adherence: Adherence models a complex patient characteristic involving affordability of medicine, psychosocial aspects of type 2 diabetes as well as "readiness to change" (Donnan, MacDonald and Morris 2002). Adherence moderates the dose of a medication and other treatments a patient effectively uses. For example, a 90% adherent patient with a 2,000 mg daily dose of Metformin will on average use 1,800 mg a day and experience blood glucose lowering effects accordingly. Depression affects patients through adherence; depressed patients take less of their medication. Adherence is independent from depression and can be low in patients who are free of depression. Seasons affect some patients' adherence; for example some patients might not follow dietary restrictions during the holidays. The strength of this effect is modeled by the *seasonal fluctuation* variable.

$$Adherence_{current} = f(Adherence_{previous}, SeasonalFluctuation, NurseEducatorEffect, DepressionEffect, RandomVariation)$$

Depression. Depression has a significant impact on the health state of a patient with type 2 diabetes (Katon, et al. 2005). If a patient is severely depressed, then their adherence is recomputed as 1 – Depression. If a patient is not severely depressed, then their adherence remains unchanged (i.e., either High or Low by description).

$$Depression_{current} = f(Depression_{previous},$$
$$Adherence*[ZoloftEffect],$$
$$PsychologistEffect, RandomVariation)$$

Responsiveness. Responsiveness (Table 2) describes individual differences between patients (i.e., inter-patient variation) with regards to how medications affect their blood glucose levels (Evans and Johnson 2001) (Wilkonson 2005). If this variable is not set, it is defaulted to 1 (average responsiveness) and will not moderate the effect of the drug on the patient. Once a patient's responsiveness profile is initialized, it does not change over the course of treatment.

Table 2. Responsiveness distributions per drug. A complete list of every drug included in the SimCare Patient Model is detailed in the next section.

Drug	Normal Distribution	Sources
Glimeperide	$\mu = 100.0\%$, $\sigma = 9.0\%$	(Campbell 1988)
Glipizide	$\mu = 100.0\%$, $\sigma = 7.0\%$	(DeFronzo 1999) (Campbell 1988)
Glyburide	$\mu = 100.0\%$, $\sigma = 7.0\%$	(DeFronzo 1999) (Campbell 1988)
Metformin	$\mu = 100.0\%$, $\sigma = 4.0\%$	(Garber, et al. 1997)
Pioglitazone	$\mu = 100.0\%$, $\sigma = 6.0\%$	(Aronoff, et al. 2000)
Rosiglitazone	$\mu = 100.0\%$, $\sigma = 10.0\%$	(Lebovitz HE, et al. 2001)
Insulin	$\mu = 100.0\%$, $\sigma = 24.0\%$	(Esposito, et al. 2008)

Pathway Variables

Pathway variables (e.g., medication effects) are updated at each clinical encounter based on the treatment moves made by a physician. The updated values of the pathway

variables are then used to compute the next patient state. The effects of the pathway variables take into account all new doses of medications, prior states of the patient and the amount of time (days) until the next patient state is computed.

For example, if a physician prescribes 500mg of Metformin to a new patient, then that patient's A1c level will be affected gradually at various future patient states (e.g., 30, 60, or 90 days later) until the full effect of the dose manifests. To maintain the accuracy of the effect that 500mg of Metformin has on the A1c variable, a time-effects curve is used to control the percent of the total effect that will eventually manifest in the patient. Furthermore, that same 500mg of Metformin will reduce SMBG levels by the same (scaled) magnitude but on a much faster time scale, thus a steeper time effects curve is used to model medication effects on SMBGs than on A1cs. Oral medications affect a daily panel (e.g., pre-breakfast, pre-lunch, pre-dinner readings) of SMBG values each to the same degree, whereas certain types of insulin can change specific SMBG values within a panel. (The full description of SMBGs and insulin is detailed in the section Noteworthy Dynamics of the SimCare Model.) Appropriate time effects curves are thus used for each pathway variable so that the patient state accurately represents the effects of all previous treatments, regardless of when the clinical encounter is scheduled.

Dose Response Curves

Dose response curves are used to compute the maximum effect any dose of a medication (or other treatment move) can have on the patient state. Dose response curves were

constructed from literature reviews of randomized clinical trial results. Each trial provided common dose amounts and corresponding intermediate health outcome. These points (e.g., 3, 4, or 5 points) were used to construct each dose response curve, with all intermediate points provided by straight-line interpolation. Figure 6 shows the dose response curve for Metformin, represented by four data points with all other values interpolated. A list of each drug included in the SimCare Patient Model is detailed in the following section Dose Response and Time Effects per Pathway Variable. A complete definition of the dose response curves (every dose value's computed effect) for every treatment is defined in Appendix A.

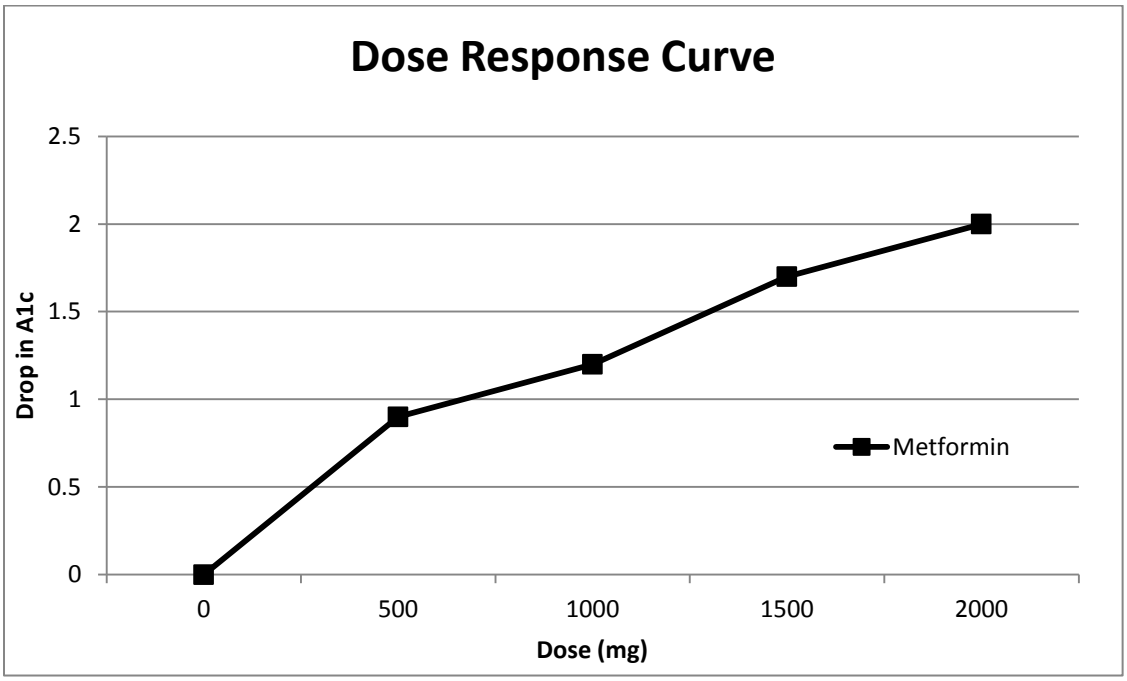


Figure 6: The dose response curve for Metformin. The complete dose response curves for all other medications can be read from Appendix A (and graphed) analogously.

Time Effect Curves

The effect of a treatment move made by a physician on a patient state changes over time (Mazze, Strock and Simonson 2001). Once a treatment move is made, the appropriate effect manifests gradually in the patient, ramping up in magnitude until it reaches the full dose effect (Figure 7).

Time effects equations update the values of pathway variables according to the amount of time the treatment effect has on each patient state variable. This models the pathophysiological effect that medications have on the patient state values over time. For example, if an effect has a "90 days to full effect" function associated with it, then after

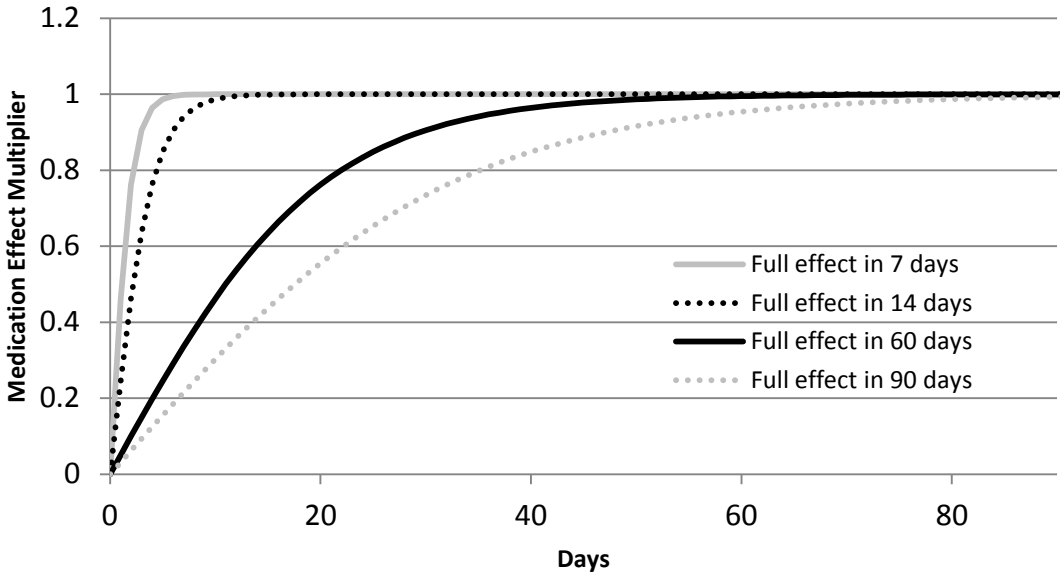


Figure 7: Sample time-effects curves from the SimCare model. Time effects curves control the time-course over which the full dose effect of each medication manifests in the patient state.

30 days, 73% of the effect has manifested. The specific equation used in the SimCare Patient model is not contained in any of the healthcare modeling literature reviewed for this thesis. Because of the difficulty of searching literature for specific equation expressions (often, healthcare modeling—even validation—literature does not explicitly state equations), it is difficult to determine if usage of this equation is novel to the SimCare Model. In any case, the important characteristics of this formula are listed and any alternative equations that satisfy these characteristics could be considered in future work.

The equation is a form of a general sigmoid function (Bishop 1995) and is based on the expression, f(x) = 1- 2/(1+e^x). There are several properties that make it appropriate for use in the SimCare Patient model, which is a domain that specifies that time is continuously moving forward (x is always >= 0):

- f(x) = 0, when x = 0
- the limit of f(x) as x pursues ∞ is 1
- f(x) is monotonically increasing, meaning clinically that a dose response could never go down at any moment in time compared to a prior moment
- x can be parameterized to adjust the severity of the slope of f(x)

The equation is stated in terms of clinical encounters and the time elapsed between two encounters:

E = encounter number, T_E = time at encounter E (in days), $T_1 = 0$

Time effect of medication moves that take full effect after *d* days made at encounter *i* at a given encounter E; *i* <= *E*.

$$TimeEffect_d(E,i) = 2\left(\frac{1}{(1+e^{\frac{T_i - T_{(E+1)}}{M}})} - \frac{1}{(1+e^{\frac{T_i - T_E}{M}})}\right);\ M = \begin{cases} .5; d = 3 \\ 1; d = 7 \\ 2; d = 14 \\ 10; d = 60 \\ 16; d = 90 \end{cases}$$

$$\sum TimeEffect_d(E,i) = 1\ if\ (T_E - T_i) \geq d$$

Dose Response and Time Effects per Pathway Variable

Tables 3-11 below list the treatment variable and the dynamics of their full pathway effects on each patient state variable. The effect of each drug is computed independently and is then added with any other affected variables in each pathway for a combined effect.

Maximum dose indicates upper bound for a medication effect; i.e. there will be no additional benefit for giving more than the indicated dose.

Maximum effect indicates the greatest possible effect that can be obtained using this drug. For example, in the best case (perfect adherence and responsiveness), titrating a patient from no Metformin to up to 2,000 mg will result in a 2% drop in A1c.

Time curve indicates the number of days until the full effect has manifested. For details, refer to the section Time Effects above.

Affected variable indicates what patient state variable is affected by the treatment. Since A1c and SMBGs are related but drug effects take longer to manifest in A1c values (A1c % is a biological pathway downstream of SMBG levels), they use different time curves. For example, increasing the dose of Metformin will lower blood glucose (as measured by SMBGs) within 14 days but it will take 90 days to see the full effect in A1c.

Table 3. Blood glucose medication effects

Treatment	Maximum Dose (mg)	Affected Variable	Days to Full Effect	Maximum Effect	Source
Metformin	2000	A1c	90	-2.0 %	(DeFronzo 1999) (Garber, et al. 1997)
		SMBGs	14	-88.8 mg/dL	(DeFronzo 1999) (Garber, et al. 1997)
Glipizide	20	A1c	90	-1.5 %	(Campbell 1988) (DeFronzo 1999)
		SMBGs	14	-66.6 mg/dL	(Campbell 1988) (DeFronzo 1999)
Glyburide	10	A1c	90	-1.5 %	(DeFronzo 1999)
		SMBGs	14	-66.6 mg/dL	(DeFronzo 1999)
Glimeperide	8	A1c	90	-1.5 %	(DeFronzo 1999)
		SMBGs	14	-66.6 mg/dL	(DeFronzo 1999)
Pioglitazone	45	A1c	90	-1.6 %	(Aronoff, et al. 2000)
		SMBGs	60	-71.0 mg/dL	(Aronoff, et al. 2000)
Rosiglitazone	8	A1c	90	-1.5 %	(Lebovitz HE, et al. 2001)
		SMBGs	60	-66.6 mg/dL	(Lebovitz HE, et al. 2001)
Glargine (Basal, Long Acting Insulin)	3units/kg /day	A1c	90	-5.6 %	(Esposito, et al. 2008)
		SMBGs	1	-248.8 mg/dL	(Esposito, et al. 2008)
Lispro (Prandial, Short Acting Insulin)	3units/kg /day	A1c	90	-2.2 %	(Rosenstock, et al. 2001)
		SMBGs	1	-97.9 mg/dL	(Rosenstock, et al. 2001)

Table 4. Blood pressure medication effects

Treatment	Maximum Dose (mg)	Affected Variable	Days to Full Effect	Maximum Effect	Source
Lisinopril	80	SBP	7	-30 mmHg	(Mogensen, et al. 2000)
	40	Creatinine	7	+0.5 mg/dL	(Mogensen, et al. 2000)
Cardizem	240	SBP	7	-20 mmHg	(Hansson, et al. 2000)
Amlodipine	10	SBP	7	-15 mmHg	(Julius, et al. 2004)
Atenolol	100	SBP	7	-15 mmHg	(Dahlof, et al. 2002)
Hydrochlorothiazide	25	SBP	7	-15 mmHg	(Materson, et al. 1993) (Carter, Ernst and Cohen 2004)
Valsartan	320	SBP	7	-30 mmHg	(Julius, et al. 2004)
	160	Creatinine	7	+0.5 mg/dL	(Julius, et al. 2004)

Table 5. Lipid medication effects

Treatment	Maximum Dose (mg)	Affected Variable	Days to Full Effect	Maximum Effect[§§§]	Source
Simvastatin	80	LDL	60	-55% mg/dL	(Jones, et al. 1998)
	40	Triglyceride	60	-25% mg/dL	(Jones, et al. 1998)
	40	HDL	60	+8 mg/dL	(Jones, et al. 1998)
Atorvastatin	80	LDL	60	-60% mg/dL	(Jones, et al. 1998)
	40	Triglyceride	60	-31% mg/dL	(Jones, et al. 1998)
	40	HDL	60	+9 mg/dL	(Jones, et al. 1998)
Gemfibrazol	1200	Triglyceride	60	-35% mg/dL	(Huttunen, et al. 1991) (Frick, et al. 1987)
		HDL	60	+15% mg/dL	(Huttunen, et al. 1991) (Frick, et al. 1987)
Fenofibrate	67	LDL	60	-5% mg/dL	(Kornitzer, et al. 1994)
	200	Triglyceride	60	-40% mg/dL	(Kornitzer, et al. 1994)
	200	HDL	60	+15 mg/dL	(Kornitzer, et al. 1994)

Table 6. Lifestyle prescriptions

Treatment	Affected Variable	Time Curve (Days)	Maximum Effect	Source
exercise	A1c	60	-0.3 %	(Rickheim, et al. 2002)
	SMBG	1	-13.32 mg/dL	(Rickheim, et al. 2002)
low fat diet	A1c	14	-0.1 %	(Rickheim, et al. 2002)
	SMBG	7	-4.44 mg/dL	(Rickheim, et al. 2002)
	Weight	60	-4 lbs	(Rickheim, et al. 2002)
high fiber diet	A1c	14	-0.1 %	(Rickheim, et al. 2002)
	SMBG	7	-4.44 mg/dL	(Rickheim, et al. 2002)
low calorie diet	A1c	14	-0.1 %	(Rickheim, et al. 2002)
	SMBG	7	-4.44 mg/dL	(Rickheim, et al. 2002)
	Weight	60	-4 lbs	(Rickheim, et al. 2002)
low sodium diet	SBP	14	-10 mmHg	(Dodson, et al. 1989)

[§§§] Lipid medications are modeled to have a percentage drop effect on LDL and triglycerides. For example, if a patient starts out with 200 mg/dL LDL, the maximum effect for Simvastatin is -55% * (200 mg/dL) = -110 mg/dL.

Table 7. Referrals

Treatment	Affected Variable	Time Curve (Days)	Maximum Effect	Source
Psychologist	Adherence	1	+10 %	(Organization n.d.)
	Depression	1	- 30 %	(Organization n.d.)
Nurse Educator	Adherence	1	+30 %	(Organization n.d.)

Table 8. Depression medication effects

Treatment	Maximum Dose (mg)	Affected Variable	Time Curve (Days)	Maximum Effect	Source
Zoloft	150	Depression	21	-90 %	(Whooley and Simon 2000)

Table 9. Disease progression

Affected Variable	6-Month Effect	Source
A1c	+0.12852%	(Levy, et al. 2004)
SMBG	+5.69988 mg/dL	(Levy, et al. 2004)
LDL	+0.1071 mg/dL	(Cowie and Harris 1995)
SBP	+2.14272 mmHg	(Cowie and Harris 1995)

Table 10. Random Variation

Patient State Variable	Normal Distribution	Sources
A1c (%)	$\mu = 0.0$, $\sigma = 0.45$	(Phillipov and Phillips 2001)
SMBG (mg/dL)	$\mu = 0.0\%$, $\sigma = 7.0\%$	(Derr, et al. 2003)
HDL (mg/dL)	$\mu = 0.0\%$, $\sigma = 7.0\%$	(Bachorik and Ross 1995)
LDL (mg/dL)	$\mu = 0.0\%$, $\sigma = 4.0\%$	(Bachorik and Ross 1995)
Triglycerides (mg/dL)	$\mu = 0.0\%$, $\sigma = 5.0\%$	(Bachorik and Ross 1995)
Systolic Blood Pressure (mmHG)	$\mu = +12.0$, $\sigma = 4.0$	(Mancia, et al. 1987), (Wen, et al. 1993)

Table 11. Weight fluctuation

Affected Variable	Days to full effect	Effect	Source
Weight	90	+/-1% lbs	(Wannamethee and Shaper 1999)

Noteworthy Dynamics of the SimCare Model

The following section details particular representations in the SimCare model for which domain experts would require greater description. These representations include the represented dynamics of blood glucose and glycosylated hemoglobin, the specific effects of insulin, the removal of medications, and combinations of multiple medications in the patient model.

Blood Glucose (SMBG) and Glycosylated Hemoglobin (A1c)

Blood glucose control is an important aspect of treating type 2 diabetes patients (the other two are blood pressure and lipids). In order to control blood glucose, physicians primarily rely on two types of measurements:

- SMBG: Self-monitored blood glucose

  A simple blood test that provides the current level of blood plasma glucose. This test is virtually instantaneous and provides a snapshot of the present glucose level in a patient's blood. These levels fluctuate throughout the day, especially after meals. Blood glucose is commonly measured in mg/dL.

- Glycosylated hemoglobin HgA1c

  A certain amount of glucose molecules present in the blood join with the hemoglobin in red blood cells. This is a one-way process, once hemoglobin has

become *glycosylated*, it will not revert. Given that red blood cells have a lifespan of about 90-120 days, the glycosylated hemoglobin level can be considered a natural moving average of blood glucose levels. A1c levels are expressed as a percentage.

In spite of the time-shifted relation between current blood glucose and A1c, researchers have identified that linear regressions can describe the relationship between these two factors (D. B. Sacks 2008). From analysis of a population of type 2 diabetes patients in the Midwestern United States, the following relationship was derived:

$$\text{Average Blood Glucose: } ABG = 44.4(A1c - 5.8) + 87$$

$$\text{or in inverse form: } A1c = (ABG-87)/44.4 + 5.8$$

In the SimCare Patient Model, both current blood glucose (BG) levels and glycosylated hemoglobin A1c are modeled (Figure 8). When a patient is initialized, the two variables satisfy the linear relationship.

At every encounter, when a treatment move is made to affect a patient's blood glucose levels, the effect is calculated on both BG and A1c. To stay true to the physiology, medication effects on A1c are delayed more than effects on BG by using different time effects curves for each medication on BG and A1c. For example, although a dose of Metformin will have the same ultimate effect on both measures, the BG effects will be fully shown in 14 days, whereas they will take 90 days to fully manifest in the A1c. The

overall dose effect of the drug is the same for A1c and SMBG; it is scaled using the slope from the regression equation (Rohlfing, et al. 2002).

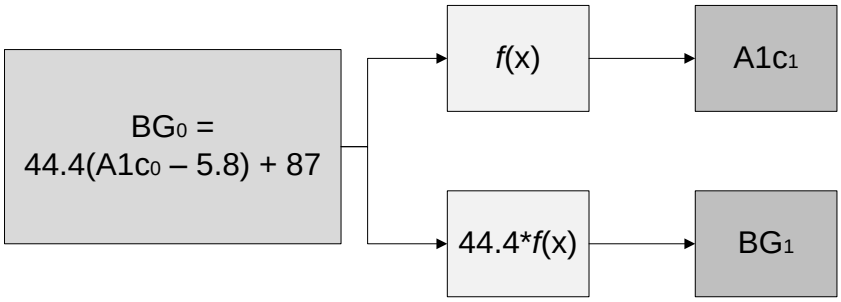


Figure 8: Conceptual model of initial relationship and separate computation of A1c and BG

In addition to maintaining the relationship between A1c and BG levels, panels of specific SMBGs are generated for the physician to use in the decision making process. These SMBG values are generated five times per day over the course of the previous three days to the current clinical encounter. The time slots are 3am, pre-breakfast, pre-lunch, pre-supper, and bed time.

Insulin Effect on Blood Glucose

Insulin is a special case for modeling because of its very fast impact on blood glucose. The model assumes that a patient receives four doses of insulin throughout the day, at breakfast, lunch, supper and bedtime.

The model makes use of two types of insulin, L and G. For every type, an impact is calculated for five times of day: 3am, breakfast, lunch, supper and bedtime. The impact is a linear combination of the doses taken throughout the day.

For example, for Glargine, the effect on blood glucose can be expressed as:

$$Insulin_L Effect_{BG} = -1244FD$$

$$F = \begin{pmatrix} .15 & .15 & .15 & .15 \\ .175 & .175 & .175 & .175 \\ .15 & .15 & .15 & .15 \\ .15 & .15 & .15 & .15 \\ .125 & .125 & .125 & .125 \end{pmatrix}, D = \begin{pmatrix} Dose_{Breakfast} \\ Dose_{Lunch} \\ Dose_{Supper} \\ Dose_{Bedtime} \end{pmatrix}$$

The dose is expressed as:

$$Dose = \frac{Units\ of\ Insulin}{Min\{Patient\ weight\ in\ kg, 136\}}$$

$$Daily\ Average\ Effect\ on\ Blood\ Glucose = \frac{\sum_{i=1}^{5}\left(\overrightarrow{FD}\right)_i}{5}$$

For the two insulin types, the dose response can be aggregated to calculate an effect on blood glucose levels for a full day (see Appendix D for details).

$$DoseResponse_{Insulin_L,BG}(x) = 73.396x$$

$$DoseResponse_{InsulinG,BG}(x) = 186.6x$$

A patient gets a linear combination of these two insulin types:

$$\alpha L + (1 - \alpha)G \;\; s.t. \alpha \in [0,1]$$

Removing Medications

Certain situations in clinical care dictate that medications that a patient is currently on need to be removed or reduced. Sometimes this is because the patient is not tolerating the medication well, other times it could be that there is a contraindication between two medications or certain health indicators rule out the use of a drug (e.g., high creatinine levels contraindicating the use of Metformin). The effects of removing or reducing medications in the patient model are an important part of the model. In most cases, these effects are simple. The dose response curve for the dose and drug are computed in the same magnitude as originally computed when the drug was added but then the direction of the effect is reversed. If 2000mg of Metformin originally decreased the patient's A1c by 2.0%, then when 2000mg of Metformin are removed from the patient, 2.0% is added back into the patient's current A1c. The time course for the removal of the drug effects to be shown is the same as for the original addition of the drug.

There are several families of drugs that affect certain aspects of the patient state in a relative rather than nominal way. For example, lipid-lowering medications reduce LDL and Triglycerides by a % change. This may cause an inaccurate replacement effect when the percent change in the opposite direction is applied to a smaller base.

Multiple Medications

Currently, the SimCare model treats the effects of individual medications independently from each other. Because of this, the effects of multiple medications are completely additive. Some research has shown that various combinations of oral antihyperglycemic medications can yield a combined effect that is less than the addition of the individual effects added together (Inzucchi 2002). This remains a possible area for improving the SimCare model.

Summary of Conceptual Validation

The SimCare model is a computational model of an individual patient receiving and responding to series of outpatient, clinical treatments. A conceptual validation of the SimCare Patient Model was conducted in order to validate the content of the model. The validation of this content is one way to show the explanatory power the model provides the user. In the case of the SimCare model, the conceptual validation accomplishes more than Sargent's criteria of making transparent the representation of the problem entity; it validates the intended mechanistic properties of the model that enable explanations and generalizations arising from emergent behavior.

The constraints of the problem were identified and their representation in the model was demonstrated. Part I of the conceptual validation organized the constraints into two constructs: the way the disease of type 2 diabetes is represented and the way outpatient clinical encounters are represented. For the representation of type 2 diabetes, sets of patient state features were described. These features belonged to the observable patient state, the unobservable patient state and the way the patient state responded to treatments. The constraints of the clinical encounter were categorized into providing information to physician decision makers, responses to physician actions, and responses to the timing of sequences of clinical encounters.

Part II of the conceptual validation showed how the constraints of the model were defined to work over time. The algorithmic theory of the model was provided as sets of patient state pathways to describe the process used to generate successive patient states. This

defined the patient state variables, the patient state pathway definitions, the patient state pathways variables and the way those variables are computed (via dose and time response curves). Additionally, a full description of the noteworthy dynamics of the SimCare model (e.g., insulin effects on the SMBG panel) was provided.

The next chapter, which contains two parts, conducts an operational validation of the SimCare Patient Model by comparing the results of simulation experiments to real observations using subjective comparisons and goodness-of-fit measures. Two contexts for the patient model to be used in simulation are controlled cohort studies and the clinical care of individual patients receiving treatments from primary care physicians. The third chapter conducts a set of experiments designed to simulate three randomized clinical trials, each with its own focus on blood sugar, blood pressure or lipids panel control. The fourth chapter conducts sets of N=1 experiments designed to simulate individual patient variation in the clinical care of a physician.

Operational Validation of the SimCare Model

The ultimate purpose of creating a computational model of a natural system is to be able to use that model to generate accurate predictions of outcomes of that system. Using such models to simulate a conceptualization of reality in order to generate predicted outcomes could range from generating or testing hypotheses about reality, answering counterfactual questions or predicting long-term outcomes yet unrealized by the natural system.

The SimCare model is a conceptual model of an individual patient with type 2 diabetes. The conceptual validation for this model has been provided in the previous chapter and enables the model to be further evaluated and generalized. Briefly, the SimCare model is a model of an individual patient with type 2 diabetes conceptualized from the perspective of a patient responding to the clinical administration of treatments by a primary care physician. This conceptualization contrasts with epidemiological or pathophysiological models. Such models approach type 2 diabetes as a population problem or a disease pathology problem, respectively, and would have different intended uses than the SimCare model.

The computational theory for the problem of an individual patient responding to physician treatment of type 2 diabetes creates intended contexts for relevant model uses—simulating individual patients and simulating groups of patients (via aggregation) responding to specific treatment protocols over time. Given SimCare's intended uses, an *operational validation* needs to be performed over these domains. As Sargent notes, an

operational validation is defined as determining that the model's output behavior has sufficient accuracy for the model's intended purpose over the domain of the model's intended applicability (Sargent 2004). Obtaining a *high* degree of confidence in a model and its results is not a product of applying a specific set of tests to determine the "correctness" of a model. Rather, the validation researchers must decide ahead of time what comparison tests are applicable for the validation experiments. Three basic approaches include the use of 1) graphs of model and system behavior data to make a subjective decision, 2) confidence intervals, and 3) hypothesis tests. It is preferable to use confidence intervals or hypothesis tests for the comparison because these allow for objective decisions. However, it is frequently not possible in practice to use either approach because a) the statistical assumptions required (often, data independence and normality) cannot be satisfied or b) there is insufficient quantity of system data available (necessitating assumptions).

In the SimCare model, a part of the patient state is conceptualized as the "unobservable" patient state meaning it is difficult or impossible to observe directly in the natural system at a given point in time. Furthermore, as is detailed in the next section, limited natural system data of cohort demographic and treatment regimen statistical properties is available in randomized clinical trial reports, making comparison difficult with the same statistical confidence intervals used in the trials. For these reasons, appropriate combinations of subjective graphical comparisons, objective confidence intervals and objective hypotheses tests are used to determine the circumstances under which the SimCare model is sufficiently accurate. Specifically, individual randomized clinical trial

simulations are compared using subjective table and graphical comparisons (due to lack of published, specified statistical data and known non-normal distributions of data), yet the model's ability to generate multiple types of predictions (SBP, HDL, A1c, etc.) over the whole series of validation trials is assessed using formal statistical measures for goodness of fit (R^2 measures). In healthcare modeling and validation of randomized clinical trial simulations, this is a common way to statistically quantify and visualize the performance of a model over a set of simulations (Eddy and Schlessinger, Validation of the Archimedes Diabetes Model 2003). It may be preferred to use direct statistical tests for intermediate outcomes comparisons; however, two limitations prevent a direct application of well-known statistical tests. Parametric tests are difficult to use because the underlying distributions of patient outcomes are known to be (highly) non-normal due to the disruptive effects of disease and physician intervention. Non-parametric tests are difficult to use because only parametric descriptions of data are published by the trials that were simulated.

Operational validation for the SimCare model is performed by conducting studies that subject simulated patients to treatment protocols taken from typical regimens of clinical care and randomized clinical trials. The first study compares the results of simulating the treatment of populations of individual, synthetic patients to treatment protocols published in three randomized clinical trials of patients with type 2 diabetes. This study demonstrates that the SimCare model can be used to conduct simulation studies for arbitrarily-sized sub-groups of patients, with randomized clinical trial simulation being a prototypical example of such a use. The three trials in this work were chosen because

they represent current treatment practices and well-accepted outcomes for each of the three main aspects of the SimCare model: 1) blood sugar intervention, 2) lipids panel intervention, and 3) blood pressure intervention. These three aspects of the SimCare model are central to representing the complex patient with type 2 diabetes. Although the focus of each trial differs, full patient state values are computed for each patient simulated in each trial, and results are aggregated for comparison. Additionally, data from these randomized clinical trials were not used in the development of the model.

Due to the nature of randomized clinical trials, detailed results are published for the focus of the trial (e.g., cohort blood sugar averages every six months throughout the trial) and additional information is shown for the pre- and post-trial follow-up (e.g., lipids levels in a blood sugar trial). The SimCare model was used to simulate these trials to the level of detail provided in the published reports. At this level of comparison, intermediate health outcomes (e.g., A1c, LDL, SBP, etc.) for the trials were compared to SimCare's predictions both in terms of cohort averages and differences between cohort means (i.e., between control and intervention). The UKPDS Risk Engine (Stevens, et al. 2001), which uses intermediate health outcome and demographic information as inputs, was used to predict patient cardiovascular event rates throughout the duration of each trial. The UKPDS Risk Engine was developed from a longitudinal study of patients with type 2 diabetes in the United Kingdom over the years 1977-1991. This particular risk engine was used in present work because it is widely recognized by the medical research community and often used by other simulation models (T. M. Group 2007). Other risk engines could be substituted in future work.

The second operational validation study (in the following chapter) compares the results of real patients receiving treatments from primary care physicians in a collaborating healthcare organization with the results of a comparable group of simulated patients receiving these same treatments. This study uses an N=1 approach to testing the SimCare model by looking at the results of individual patient simulations (Guyatt, Sackett and Taylor, et al. 1986). This approach is called "N of 1" because it narrows the scope of each experiment to the comparison of an individual patient to an individual simulated patient. The second study of operational validation assesses the ability of the SimCare model to represent meaningful variation over time for individual patients receiving typical forms of clinical care in the real world. The types of patients and treatments in randomized clinical trials are tightly controlled both for aspects of patients (e.g., patients must demonstrate high adherence to treatments) and for specific treatments (e.g., the intervention and control regimens). However, in the real clinical environment, primary care physicians are faced with a wide variety of patients (e.g., non-adherent), treatment scenarios and patient reactions to prior treatments. This makes the N=1 arm of operational validation important because it demonstrates that the SimCare model can be used to represent individual patients responding to clinical treatments over time, as well as controlled populations of patients receiving fixed regimens of care.

The framework of operational validation is used to test the accuracy of predictions of the model. This complements the conceptual validation conducted in reference to the mechanistic, explanatory power of the model. In order to test the accuracy of the model's

predictions, both subjective comparisons of data and goodness-of-fit measures are conducted. In the operational validation, two types of studies are being conducted to test two different uses of the model, and different goodness-of-fit measures are developed for each test. One method tests predictions in the context of randomized clinical trials; the other tests the context of replicating individual patients receiving typical clinical treatments.

Three aspects of the randomized clinical trial simulations are tested. They are the ability of the SimCare model to 1) simulate intermediate health outcomes for independent cohorts of patients as reported in the trial, 2) simulate proportional differences between intermediate health outcomes of relevant cohorts (e.g., relative risk), and 3) generate risk data using a third-party risk engine compared to the cardiovascular disease (CVD) event rates for each trial. For each aspect listed, R^2 measures are used to test the goodness-of-fit of the model predictions to the observed results of the trials.

For the individual patient study, a novel approach to testing individual patient variation over time was created through a series of N=1 experiments. The purpose of this is twofold: 1) demonstrate that SimCare predicts a variety of responses across individual patients receiving **typical forms of clinical treatments**, and 2) test that SimCare can predict **individual variation** in patients who receive these typical forms of treatment. This study examines each patient throughout a typical range of responses by establishing a "best" and "worst" case bound on that patient's treatment path. "Typical" forms of treatment are defined by a set of treatments representing expert guidelines (Nathan, et al.

2009), and typical responses to treatments are defined as statistical standard deviations away from responses of patients with average characteristics. If the observed patient outcome is within these simulation bounds, then the N=1 experiment for that patient is deemed a success. If the observation exceeds the bounds, then the experiment is deemed an error. Analysis of the occurrence rate of errors throughout the model domain space exposes potential bias in the SimCare model.

Randomized Clinical Trial Validation

In the first study of the operational validation of the SimCare model, randomized clinical trials are used to test predictions of intermediate health outcomes by applying standard forms of treatments to cohorts of patients on an individual patient basis (i.e., clinical encounter by clinical encounter). Three benchmark clinical trials were selected to address each of the main three patient state variables of the SimCare model: HgbA1c, Lipids, and Blood Pressure. The trials chosen for this validation exercise were ADVANCE Blood Glucose, CARDS (Lipids), and ADVANCE Blood Pressure. These trials were chosen because they are widely accepted and were conducted recently enough to include current treatment practices and medications.[****] None of these trials was used to build the model, and each will be described in detail in its own section.

There is a considerable amount of variety in how a given randomized clinical trial is conducted and reported. For example, if a trial is testing the effects of a lipid-lowering medication, then detailed information about the cohorts' lipids levels is published, often (not always) at six month intervals throughout the trial. Additional information, such as the A1c and blood pressures, is usually published for the beginning of the trial and the end of the follow up period in the form of an average and a standard deviation. For these additional measures, no intermediate information is published. Most trials also publish

[****] This does not represent a complete list of all randomized clinical trials the SimCare model could simulate. Additional trials such as HPS (H. P. Group 2003), UKPDS (Holman, Paul, et al., 10-Year Follow-up of Intensive Glucose Control in Type 2 Diabetes 2008) (Holman, Paul, et al., Long-Term Follow-up after Tight Control of Blood Pressure in Type 2 Diabetes 2008), ACCORD (T. A. Group 2008), Steno (Gaede, Vedel, et al. 1999), and ALLHAT (Rahman, et al. 2005) have been considered for this work but were excluded for various specific reasons ranging from antiquated treatments to controversial results. These trials could be simulated in future validation work.

some form of macrovascular (e.g., heart attacks and strokes) and microvascular (e.g., blindness and amputation) event rates throughout the trial, though the level of event category description varies (e.g., all macrovascular events vs. myocardial infarctions). Because of this variety in reporting trial information, simulated populations of patients are generated to match real populations of patients on a trial by trial basis. The reporting of simulated results for each trial is also conducted on a trial by trial basis, with the main focus of each analysis comparing relevant intermediate health outcomes from the focus of the trial with the relevant observable features predicted by the model (e.g., analyzing blood pressure comparisons from a blood pressure-focused trial).

Generating a synthetic cohort for simulated treatment can be accomplished in a variety of ways (T. M. Group 2007). For example, if a cohort has an average A1c = 8.4%, each patient in the cohort could be assigned an A1c value of 8.4%. Although this approach maintains the reported average, it is not descriptive of the underlying variability of inter-patient A1c levels throughout the cohort. To address this issue, the following method was used. First, a super-population of 10,000 synthetic patients was generated to represent sampled, real patient populations with type 2 diabetes. The real patient source data was gathered by a collaborating healthcare organization operating clinics throughout Minnesota and consisted solely of patients with type 2 diabetes. This enabled the use of a data set that was descriptive of a large number of type 2 diabetes patients over the variables used as inputs by the SimCare model, such that the relationships among these variables could be derived in a way that was not available in medical research literature. However, this approach was also a limiting factor as the local population being treated by

the Minnesota healthcare organization may not be fully representative of populations of patients throughout the United States and the world. Such background discrepancies across population-based studies (and models) is a subject that will be touched upon often in this thesis and remains a constant source of possible error in healthcare research.

In constructing the distributions to describe observed patient data, a separate distribution was formed to match each relevant patient feature (e.g., A1c, LDL, etc). Furthermore, correlations among distributions in the real patient data were computed and used to tie separate distributions together, so that if statistical relationships existed between patient features in the real population, they became part of the synthetic patient super-population. From these correlated distributions, a super-population of 10,000 patients was generated.

To create specific cohorts that matched those described in the clinical trials, a sampling procedure was used. The sampling procedure established a descriptive population goal using the group health statistics of each real cohort at the beginning of the published trial. Typically, this included the A1c, SBP, LDL, HDL, and Triglycerides of each group. Each health characteristic was equally weighted in importance for generating a cohort that matched the goal statistics. A simulated patient was drawn from the super-population and considered for inclusion in the simulated cohort. If the addition of that patient brought the cohort health statistics further away from the goal values, then that patient was discarded. If that patient brought the synthetic cohort statistics closer to the goal statistics, then that patient was added to the cohort. This sampling process continued until the simulated cohorts matched the number of patients in each real cohort, the

average health variable values, and the inter-patient variability of the patient distributions in each cohort. If a measure of variability was not published by the trial, then the existing variability in the modeled super-population was used.

Another aspect of simulating randomized clinical trials is simulating the treatments given by the physicians throughout the trial. Although the specific treatments in question are tightly described and controlled (intervention vs. placebo), any additional treatments are less well-specified. In order to simulate the types of treatments a physician participating in a randomized clinical trial would administer to a given patient, the drug formulary and encounter frequency described by the trial were identified. This simulated treatment regimen used the given medication formulary and scheduling regimen in order to treat every feature of each patient one encounter at a time over the course of the simulation.

Although each clinical trial simulation is unique, the general method followed was similar:

1. Generate cohorts of individual synthetic patients that match each initial cohort statistical description in the trial.
2. Generate a pharmaceutical formulary for treating patients and scheduling encounters in the placebo and intervention cohorts based on the formularies stated in the published trial protocol.
3. Simulate the trial for the stated duration of time using the formulary and dosing schedule as defined in the trial to generate representative treatments per clinical encounter to each patient.
4. Extract comparative statistics of results from the database of simulated treatments.

5. Generate graph and table information for comparison with each trial based on the focus and availability of information published.
6. Generate cardiovascular disease (CVD) event rates over the course of the trial using SimCare patient treatment data as input for a third-party risk predictor.

The SimCare model does not explicitly generate macrovascular events, so the UKPDS Risk Engine was used to simulate the risk of major cardiovascular events for individual patients (Stevens, et al. 2001). The UKPDS Risk Engine was derived based on a population study of patients with type 2 diabetes in the United Kingdom. Self-reported ethnic variables of White-Caucasian, Afro-Caribbean, and Asian-Indian were included as input variables for stratification of risk factors. However, other population-based variables (e.g., current healthcare standards, diet, etc.) were not included in the model and remain a possible source of error when using the UKPDS Risk Engine on different patient populations.

At yearly time points (e.g., 0, 1, 2, 3, and 4 in a 5 year trial) the UKPDS Risk Engine inputs were taken from individual patient records generated by SimCare, and a probability for an event was estimated for the following year for each patient. Using a Monte Carlo method (James 1980), random numbers were drawn for each patient and cardiovascular events were assigned to have happened or not. Monte Carlo methods are a useful way to generate observations (e.g., events) based on probabilities (e.g., risks). Since probabilities are supplied by the UKPDS Risk Engine, but events are not directly simulated, the Monte Carlo method simulated the discrete occurrence of events. This was a necessary part of the simulation as patients who experience such events in the

actual trials are immediately removed from the trial. However, Monte Carlo methods are susceptible to a certain kind of error, in that a patient with a risk level of 1% might have a heart attack in a single simulation. If such a low-probability event were to happen too many times in a single simulation (due to randomness), the resulting event rate would not accurately represent the underlying risks of the group. For this reason, Monte Carlo simulations must be repeated many times using the same risk data so the average simulated event rate represents the underlying patient risk levels.

In keeping with trial protocols, if an event (e.g., non-fatal stroke) happened to a simulated patient, this patient was then withdrawn from the study (i.e., they were not included in the remaining event simulations for that trial). After the Monte Carlo event simulation was carried out for each year of treatment, the events were totaled and divided over the original patient population baselines to compute an average total event rate for each cohort in each trial. To lower the first-degree sampling error associated with Monte Carlo methods, the simulation was independently conducted 100 times with the same risk factor data. The average total event rate for each event category was then compared to the published trial total event rates.

As mentioned previously three randomized clinical trials were chosen for simulation because they represent current treatment practices and available drugs as well as well-accepted outcomes in each of the three main patient state aspects of the SimCare Patient Model. The simulation for each trial is reported separately and details the ability of the SimCare model to represent health outcomes for Hemoglobin A1c (A1c), lipids (LDL,

HDL, Triglycerides), and blood pressure as reported in trial data. The trials consisted of patients with type 2 diabetes, and none was used to build the SimCare model. Results of each simulation are shown in table and graph form similar to those shown in the published reports. A discussion of the results is conducted on a trial by trial basis. An overall analysis of the three aims of the study is conducted using correlations to test SimCare's ability to predict 1) intermediate health outcomes for each cohort, 2) relative differences between the control and intervention groups, and 3) CVD event rates using a third-party risk predictor.

ADVANCE – Blood Glucose

The Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation (ADVANCE) study contained two arms to study the disjoint effects of intensive blood glucose and blood pressure control in diabetic patients (T. A. Group 2008). For the purposes of the SimCare simulation, the ADVANCE study will be treated as two separate studies, ADVANCE Blood Glucose (ADVANCE-BG) and ADVANCE blood pressure (ADVANCE-BP).

The ADVANCE-BG clinical trial was intended to show the effects of intensive blood glucose control on macrovascular and microvascular events. The major findings of the paper showed a significant difference in cohort A1c levels after 5 years of treatment (7.3% vs. 6.5%) but no decrease in macrovascular event rates. The combined microvascular and macrovascular event rate showed a 10% relative risk reduction in the intervention group, primarily due to a 21% relative risk reduction in nephropathy (kidney disease). The trial found no link between A1c differences at this level and a macrovascular event rate reduction. This study simulates the trial as 5 years of treatments to predict A1c values. Additional intermediate health outcome variables are also reported as available.

Method

Having been carried out in 20 countries, over several continents, and using 215 collaborating centers, the ADVANCE-BG trial is a difficult trial to simulate accurately.

Table 12 shows the initial cohort characteristics in both the ADVANCE-BG trial and the SimCare simulated populations. Patients were randomly drawn from the original 10,000 patient super-population described in the introduction until the average health characteristics matched those in ADVANCE-BG.

Table 12. The initial parameter settings for each cohort in the actual and simulated populations for the blood glucose arm of the ADVANCE trial simulation. The main focus of the trial is shown in bold.

Initial Population Characteristics	ADVANCE-BG		SimCare	
	Intensive	Control	Intensive	Control
N	5571	5569	5571	5569
% Female	42.60%	42.30%	43.90%	41.40%
Age at Diagnosis (sd)	58(9)	58(9)	58(9)	58(9)
Duration of Disease(sd)	7.9(6.3)	8.0(6.4)	8.0(6.3)	7.9(6.3)
A1c(sd) -- %	**7.48(1.65)**	**7.48(1.63)**	**7.52(1.55)**	**7.51(1.47)**
SBP(sd) -- mmHg	145.0(21.7)	145.0(21.4)	145.7(19.6)	145.4(19.5)
LDL(sd) -- mg/dL	120.7(40.2)	120.4(39.5)	120.7(34.8)	120.4(34.8)
HDL(sd) -- mg/dL	48.8(13.5)	48.4(13.5)	48.8(12.8)	48.8(12.4)
Median Triglyceride -- mg/dL*	141.6	145.1	157.5	159.3
Adherence	High	High	100%	100%

*Values of triglycerides in the simulated population had a higher median value due to ethnic disparities between ADVANCE populations and U.S. populations. For example, the median triglyceride values reported for SimCare are generally similar to those reported in other studies conducted in the U.S. and Canada (Sacks, et al. 1996) (T. A. Group 2008), generally higher than those conducted in Asia, and generally lower than those conducted in Scandinavian countries (H. P. Group 2003) (Gaede, Vedel, et al. 1999). Because no detailed information about the distribution was reported by the ADVANCE group, the original super-population distribution was used in the simulation.

A drug formulary and dosing schedule were defined in order to represent the treatments administered to patients under the control of the trial regimen. To match the scheduling intensities in ADVANCE-BG, patients in the control group made clinical visits two months apart for the first six months and then every six months after that (Table 13). Patients in the intensive group made clinical visits every month for the first six months and then every three months after that. These scheduling differences as well as the specific pattern of scheduling highlight the ability of the SimCare model to simulate specific clinical encounters between individual patients and physicians over time.

Table 13: ADVANCE-BG drug formulary available for blood glucose, blood pressure, and lipids panel management. Maximum doses are included in parentheses where appropriate.

Treatment Formulary	ADVANCE-BG		SimCare	
	Intensive	Standard	Intensive	Standard
Medication (max dose)				
Blood Glucose	Gliclazide		Glipizide (20mg)	
	Metformin		Metformin (2000mg)	
	Thiazolidinediones		Pioglitazone (45mg)	
	Basal Insulin		Basal Insulin	
	Prandial Insulin		Prandial Insulin	
Blood Pressure	Up to Local Physician*		Lisinopril (80mg)	
			Cardizem (240mg)	
Lipids Panel	Up to Local Physician*		Simvastatin (80mg)	
Scheduling				
First Six Months	30 Days	60 Days	30 Days	60 Days
After 6 Months	90 Days	180 Days	90 Days	180 Days
Goal				
A1c (Relapse Trigger)	Up to Local Physician*		6.0%(6.5%)	7.0%(7.0%)
LDL			105	
SBP			135	

*The ADVANCE-BG study did not publish these details. In order to simulate the trial, assumptions or substitutions for each area were made. For medications, collaborating physicians suggested viable substitute drugs within the same families of listed medications if the explicit drug used in the actual trial was not currently being modeled by SimCare. Evidence-based goals were derived empirically using trial and error.

Because the ADVANCE-BG trial was conducted over such a wide set of countries and clinics, each with different practice standards and available medical formularies, a specific set of drugs was not published as the overall trial standard. The formulary in Table 13 was used to treat all patients in both groups and was chosen in consultation with collaborating physicians to best represent the trial. The main difference between the ADVANCE-BG intensive and control groups is in scheduling intensity (i.e., not in the presence of a specific drug vs. placebo).

The basic algorithm for treating a given patient at each encounter was that if a patient was above the evidence-based goal for a given health characteristic, then an additional treatment move (e.g., titrate current medication or start a new medication) was made. In this way, the goals were used to trigger additional treatments, if needed, for each encounter, or hold treatments constant if the goal had already been reached. Since evidence-based goals are not published parts of clinical trials, they were derived empirically by running multiple simulations using differing goals and selecting the goal for each feature that generated the most accurate results. The goals used for this simulation were A1c <= 6.0% and 7.0% for the intensive and conventional groups, respectively. Additional goals of LDL <= 105 mg/dL and SBP <= 135 mmHg were used for both groups. Table 14 shows the intermediate health outcomes for all four cohorts after five years of treatment.

Table 14: Intermediate health outcomes published from the ADVANCE-BG clinical trial and simulated by the SimCare model. Results were tabulated after 5 years of treatment.

End of Follow-up	ADVANCE-BG		SimCare	
	Intensive	Standard	Intensive	Standard
N	5571	5569	5571	5569
A1c(s.d.)-%	**6.53(.91)**	**7.30(1.26)**	**6.56(.85)**	**7.35(1.20)**
SBP(s.d.)-mm Hg	135.5(17.6)	137.9(18.4)	136.5(10.3)	136.0(10.6)
LDL(s.d.)-mg/dl	102.2(38.5)	102.6(41.0)	103.7(15.5)	103.3(15.5)
HDL(s.d.)-mg/dl	48.4(13.5)	48.4(13.5)	52.2(13.9)	51.5(13.2)
Median Triglyceride-mg/dl	128.3	140.7	123.0	123.9

The main focus of the ADVANCE-BG simulation is to test SimCare's ability to predict cohort A1c levels. The additional reporting of lipids and blood pressure values is for the purpose of full disclosure (as it is also done in the published report for the trial). More

detailed data for blood glucose levels is reported in six-month increments for each cohort in Figure 9.

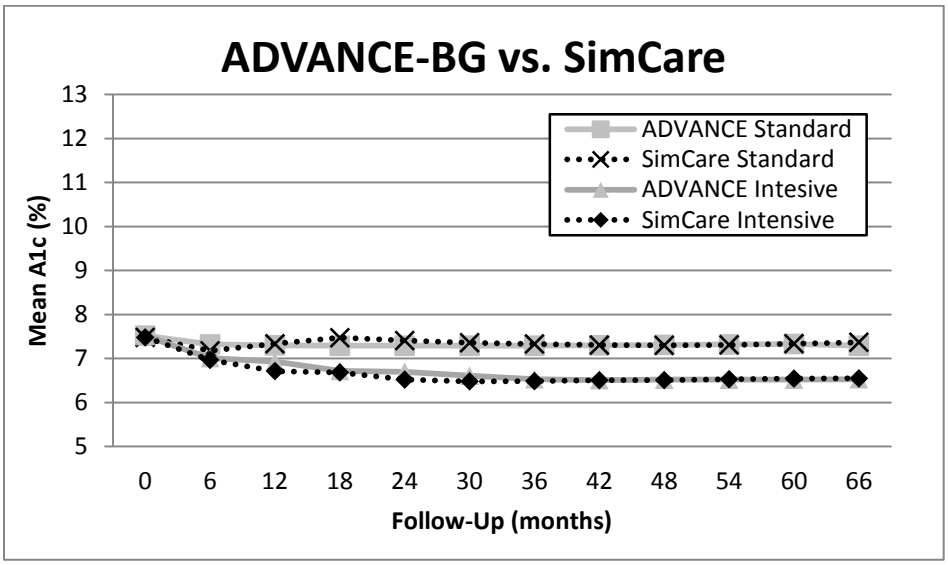


Figure 9: Mean A1c levels for each cohort (intensive and standard) in the ADVANCE-BG and SimCare simulation over the five and a half years reported in the published trial.

Discussion

The ADVANCE-BG trial represented both an important and difficult trial to simulate. The difficulty of simulation stemmed from the heterogeneity inherent in a multi-country trial taking place in over 200 clinics, each with different practices and available medical tools. Populations in different regions of the world with type 2 diabetes can vary in certain baseline characteristics, especially triglycerides. It is this difficulty that shows the importance of simulating the trial. As it turns out, only basic simplifying assumptions were needed to run the simulation (e.g., evidence based goals, population characteristics).

As was shown in Table 14, the follow-up A1c values of the simulation matched actual values in both mean and standard deviation. In addition to blood glucose, simulated systolic blood pressure means also closely matched real means, although standard deviations were smaller for the simulated cohorts (41.5% and 42.4% lower, intensive and standard cohorts, respectively). This over-treatment of population variation is an example of the difficulty in accurately simulating specific treatments on an encounter by encounter basis from published clinical trial data. The defined protocols used to treat blood glucose, blood pressure and lipids levels throughout the simulated trial did so with equal priority at every encounter. The result was that for some simulated patients, medications were increased for all three patient state variables if all three were above the evidence-based goal. In clinical trials, increasing multiple medications at the same encounter is less likely to occur due to trial physicians and primary care physicians not always being the same for each patient. Because primary care physician treatment protocols are not published as trial data, the overall treatment protocol was simulated. This assumption over-treated some simulated patients in this context with initial high blood pressure more often than they were treated in the trial. A similar phenomenon is seen in the LDL results where many patients were brought to the evidence-based goal but patients with high initial LDL were over-treated throughout the trial (in some cases at every simulated encounter), thus reducing cohort variation in the simulation.

The over-use of LDL-lowering treatments available to the SimCare model also led to a higher than observed simulated HDL mean (raising HDL is a common effect of many LDL-lowering statins). The differential between triglyceride outcome means in the real

patient cohorts (128.3 vs. 140.7, intensive and standard cohorts, respectively) was not explained in the published report, nor was it reproduced in simulated results (simulated cohorts resulted in similar triglyceride levels). Therefore, it is difficult to know how to characterize the lipids portion of this simulation experiment.

By using the UKPDS Risk Engine to simulate events, total event rates were reported based on intermediate health outcomes from the SimCare simulation. The results can be seen in Table 15. Nonfatal Coronary Heart Disease (CHD) and Death from cardiovascular events show nearly identical event rates between simulated and actual results both in terms of comparison across cohorts and differences between cohorts. Nonfatal Stroke events were nominally under-represented by the SimCare/UKPDS models, although the difference between the cohorts within the trial was maintained. A major cause of strokes is high blood pressure. As was mentioned previously, the SimCare model simulated protocols in a way that over-treated patients with high blood pressure (during the course of their primary treatment of blood glucose), resulting in smaller variation in distribution of SBP values in both cohorts (i.e., fewer patients had high SBP throughout the simulation). This explains the under-representation of stroke rates in the simulated population.

Table 15: Using the UKPDS Risk Engine, primary end points (CVD events) were simulated for the ADVANCE-BG trial using SimCare's predicted intermediate health outcomes at each yearly checkpoint.

Primary End Points - # of Total Events(%)	ADVANCE-BG		SimCare	
	Intensive	Standard	Intensive	Standard
Nonfatal Stroke	214(3.8)	209(3.8)	168(3.0)	165(3.0)
Nonfatal CHD	153(2.7)	156(2.8)	154(2.8)	162(2.9)
Death from Cardiovascular events	253(4.5)	289(5.2)	268(4.8)	301(5.4)

CARDS – Lipids

The Collaborative Atorvastatin Diabetes Study (CARDS) randomized clinical trial was designed to study the lipid-lowering effects of 10mg Atorvastatin daily on populations of patients in the U.K. and Ireland with type 2 diabetes (Colhoun, et al. 2004). The primary reason to control lipids in such populations was to decrease the occurrence of major cardiovascular events. This trial was also used as a validation exercise for several clinical models of type 2 diabetes at the Fourth Mt. Hood Challenge (T. M. Group 2007). CARDS is a unique trial to simulate because it offers the ability to test the accuracy of lipids predictions from SimCare with both the CARDS populations and the predictions of several other type 2 diabetes models published in the fourth Mt. Hood Challenge.

The purpose of the CARDS simulation is twofold: 1) compare the intermediate health outcomes for two cohorts (intervention and control) over the 4-year time period in the study with the risk factors generated by the SimCare model, and 2) compare the total rates for CVD events in the trial with the event rate predictions of SimCare and the other models at the Mt. Hood Challenge.

Method

Simulated patient populations were created to match the characteristics of the control and intervention groups in the original CARDS trial. These simulated populations matched CARDS populations in the mean values, standard deviations, and proportions (Table 16), including number in each group ($N_{placebo}$ = 1410, $N_{intervention}$ = 1428), age (mean = 61),

gender (women = 32%), ethnicity (white = 95%), LDL (mean = 117), HDL (mean = 54), Triglycerides (mean = 149), Smoking status (smoking = 23%), SBP (mean = 144), and HgbA1c (7.9%). CARDS did not publish correlations among these variables, so correlations from the super-population of synthetic patients were used when constructing the trial cohorts.

Table 16. The initial cohort and treatment statistics for the CARDS experiment using the SimCare model. The focus of the trial is shown in bold.

Initial Population Characteristics	CARDS Study		SimCare Study	
	Placebo Group	Atorvastatin Group	Placebo Group	Atorvastatin Group
N	1410	1428	1410	1428
Age				
<60	529	558	529	558
60-70	708	703	708	703
>70	173	167	173	167
Demographics				
Women	453	456	453	456
White	1326	1350	1326	1350
Smokers	323	308	323	308
Initial Health Characteristics				
LDL(sd)-mg/dl	**117(27)**	**118(28)**	**117(25)**	**118(25)**
HDL(sd)-mg/dl	**55(13)**	**54(12)**	**55(10)**	**54(12)**
Med Trig (Range)-mg/dl	**148(104-212)**	**150(106-212)**	**158(94-240)**	**164(101-240)**
SBP(sd)	144(16.1)	144(15.9)	144(14.1)	144(14.3)
Mean HgbA1c	7.81(1.39)	7.87(1.42)	7.75(1.6)	7.77(1.62)
Adherence	High	High	100%	100%

In order to simulate the treatment conditions of CARDS without fully being able to simulate every aspect of the trial, the following assumptions were made about the patient populations being treated (Table 17): 1) patients in the control group were not given additional lipid-lowering medication, 2) patients in the intervention group were given first 10mg Atorvastatin daily and up to 40mg Simvastatin if evidence-based goals for

lipids were not met, 3) conventional treatments for hypertension and blood glucose control were administered to both groups using evidence-based treatment protocols (e.g., treat-to-goal) and defined drug dose formularies. Following an additional parameter of the CARDS study, simulated patients were seen on a monthly basis for the first 3 months and on a semi-annual basis thereafter. Simulated patients were treated under these conditions for four years. At the end of the follow-up period, intermediate health outcome results were reported for each cohort.

Table 17. The treatment and scheduling experiment design for the CARDS simulation using the SimCare model.

<table>
<tr><td rowspan="2">Treatment Formulary</td><td colspan="2">CARDS</td><td colspan="2">SimCare</td></tr>
<tr><td>Placebo</td><td>Atorvastatin</td><td>Placebo</td><td>Atorvastatin</td></tr>
<tr><td colspan="5">Medication (max dose)</td></tr>
<tr><td rowspan="5">Blood Glucose</td><td colspan="2" rowspan="5">Up to Local Physician*</td><td colspan="2">Metformin (2000mg)</td></tr>
<tr><td colspan="2">Glipizide (20mg)</td></tr>
<tr><td colspan="2">Pioglitazone (45mg)</td></tr>
<tr><td colspan="2">Basal Insulin</td></tr>
<tr><td colspan="2">Prandial Insulin</td></tr>
<tr><td rowspan="2">Blood Pressure</td><td colspan="2" rowspan="2">Up to Local Physician*</td><td colspan="2">Lisinopril (40 mg)</td></tr>
<tr><td colspan="2">Cardizem (240 mg)</td></tr>
<tr><td rowspan="2">Lipids Panel</td><td>Atorvastatin (0 mg)</td><td>Atorvastatin (10 mg)</td><td>Atorvastatin (0 mg)</td><td>Atorvastatin (10 mg)</td></tr>
<tr><td colspan="2">Additional Statin up to Local Physician*</td><td>NONE</td><td>Simvastatin (40 mg)</td></tr>
<tr><td colspan="5">Scheduling</td></tr>
<tr><td>First Six Months</td><td colspan="2">30 days</td><td colspan="2">30 days</td></tr>
<tr><td>After Six Months</td><td colspan="2">180 days</td><td colspan="2">180 days</td></tr>
<tr><td colspan="5">Goal</td></tr>
<tr><td>A1c (Trigger)</td><td colspan="2" rowspan="3">Up to Local Physician*</td><td colspan="2">8.0% (8.3%)</td></tr>
<tr><td>LDL</td><td colspan="2">100 mg/dl</td></tr>
<tr><td>SBP</td><td colspan="2">145 mg/dl</td></tr>
</table>

*The CARDS study did not publish these details. In order to simulate the trial, assumptions or substitutions for each of these areas were made. For medications, collaborating physicians suggested viable substitute drugs within the same families of listed medications if the explicit drug used in the actual trial was not modeled by SimCare. Evidence-based goals were derived empirically using trial and error.

Results

Table 18 shows the CARDS results alongside SimCare's simulated results. SimCare tended to reduce LDL and Triglyceride levels and increase HDL levels slightly more than in CARDS, while at the same time reducing inter-patient variation. Mean A1c and SBP levels were controlled closely to CARDS, though variation in these feature distributions was lower.

Table 18: Intermediate health outcomes published from the CARDS clinical trial and simulated by the SimCare model. Results were tabulated after four years of treatment.

End of Follow-up	CARDS		SimCare	
	Placebo	Atorvastatin	Placebo	Atorvastatin
N	1410	1428	1410	1428
Total Cholesterol(s.d.)-mmol/L*	**4.73 (0.91)**	**3.69 (0.84)**	**4.81 (1.04)**	**3.76 (0.86)**
LDL(s.d.)-mmol/L	**3.12 (0.80)**	**2.11 (0.70)**	**3.03 (0.63)**	**1.87 (0.48)**
HDL(s.d.)-mmol/L	**1.23 (0.30)**	**1.26 (0.30)**	**1.41 (0.24)**	**1.58 (0.27)**
Triglyceride(s.d.)-mmol/L	**1.90 (1.10)**	**1.61 (0.93)**	**1.84 (0.85)**	**1.56 (.57)**
HgbA1c(sd)-%	8.1 (1.5)	8.3 (1.5)	8.23 (1.1)	8.25 (1.1)
SBP(sd)-mm Hg	144 (17)	143 (17)	144 (8)	143 (8)

*total cholesterol is not explicitly simulated by the SimCare model but is computed based on the following equation: tc = (trig/5) + ldl + hdl.

The main focus of the CARDS trial was on controlling lipids levels. Detailed cohort information for each lipid measure was published. At each yearly time point in the simulation, intermediate health outcome data (e.g., LDL, HDL) were gathered for each patient, such that each patient had data listed for years 0, 1, 2, 3, and 4. Average intermediate health outcomes for each group were tabulated and graphed for comparison to graphs published in CARDS as shown in Figure 10.

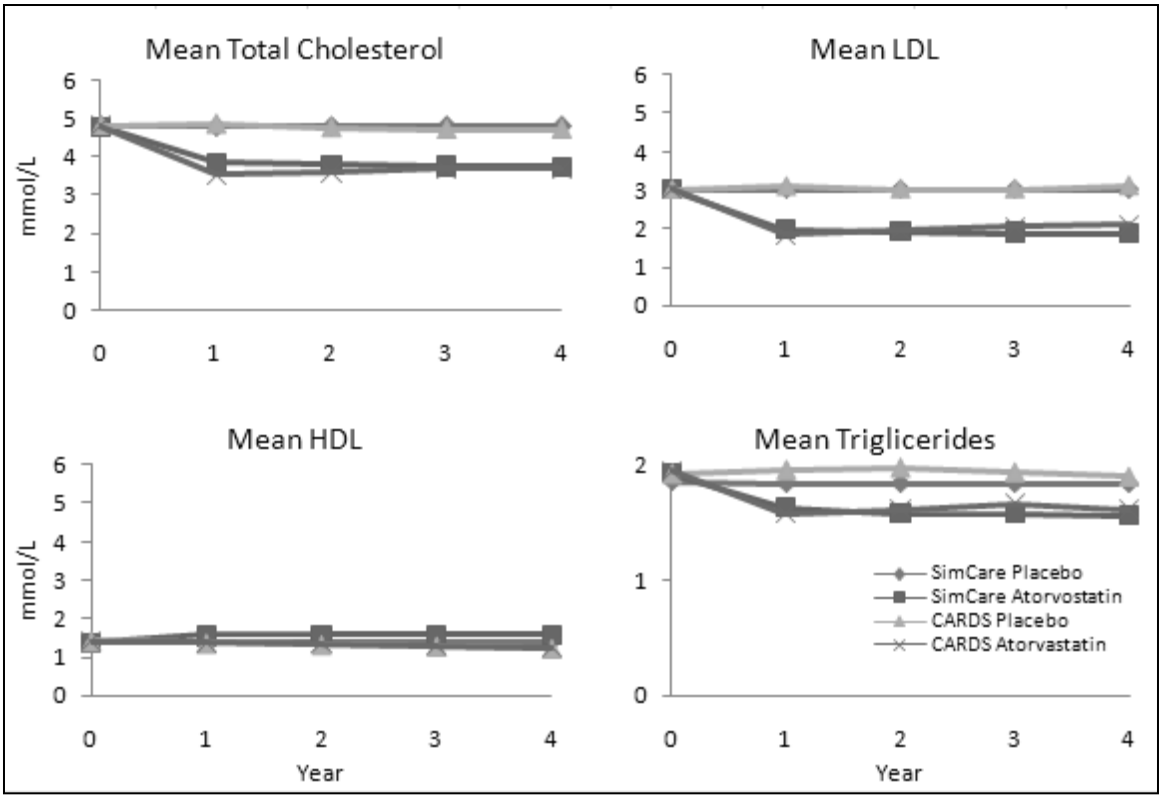


Figure 10: Graphs of lipid levels over the course of four years of treatment. To convert from mg/dL (SimCare units) to mmol/L (CARDS units), a multiplier of 38.7 was used for LDL and HDL; 88.5 for triglycerides.

Discussion

One of the interesting aspects of the CARDS simulation is that not only can the UKPDS Risk Engine be used to compare event rates to the published trial's event rates, but predicted event rates can also be compared to predicted event rates from several other "clinical models" of type 2 diabetes. As can be seen in Table 19, the risk models from the Mt. Hood Challenge tended to underestimate the actual event rates, as well as the difference between the control and intervention cohorts. The modelers participating in the challenge attributed differences in prediction results to their different underlying assumptions of CARDS populations (e.g., in initial distributions and correlations) as well as the different computational theories of each of their models (e.g., some models

included specific treatments, others did not). The Sheffield model exceeded the other models in event rate prediction levels but had a lower cohort difference prediction. The other models in general exhibited an under-prediction bias for difference in event rates between the control and intervention cohorts. (See Appendix F for a review of the computational models in the fourth Mt. Hood Challenge.)

Table 19: Three aspects of simulated CVD events for each model as reported in the fourth Mt. Hood Challenge. The first two columns show each model's predictions for cohorts of the CARDS trial. The third column shows the difference between each model's predictions for the cohorts.

Acute CVD Event Rate	Control	Intervention	Difference
CARDS	13.4	9.4	4.0
SimCare	12.9	9.7	3.2
CDC/RTI	10.2	6.9	3.3
EAGLE	8.4	0	N/A
Cardiff	9.2	6.7	2.5
Sheffield	12.4	9.6	2.8
UKPDS Group	10.4	7.4	3.0

The SimCare model predicted the risk factors for the CARDS study for each lipids panel intermediate health outcome published in the CARDS clinical trial. The data were representative of the CARDS data in the important ways of 1) appropriate separation of control and intervention groups within the first year, and 2) accurately representing resulting average values for both groups throughout the 4-year trial.

Using the UKPDS Risk Engine to predict CVD events resulted in a comparable 4-year, total event rate to the CARDS study in several important ways (Figure 11). The main aspects of the event rates were the control event rate, the intervention event rate, and the difference between the two. Two groups of data in the figure are apparent. The first

group is the CARDS trial, the SimCare model and the Sheffield model. The other group is the UKPDS Group, CDC/RTI model and the Cardiff model. (The EAGLE model was included for completeness's sake but did not simulate the CARDS intervention group).

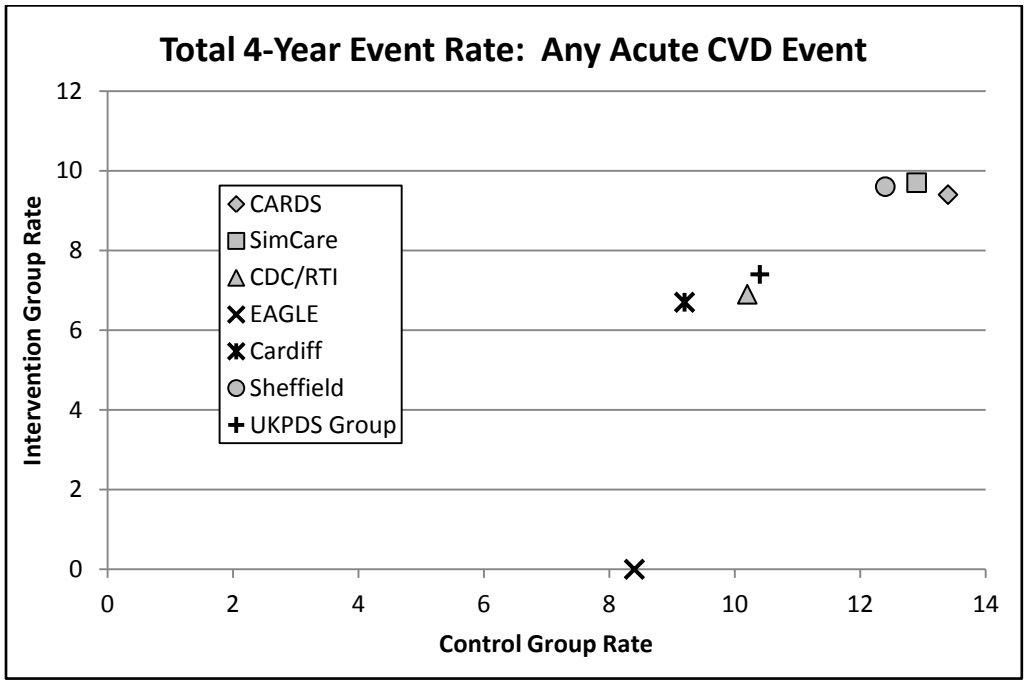


Figure 11: This graph shows the cardiovascular disease event rate for five clinical models from the Mt. Hood Challenge, SimCare and the actual event rate from CARDS. Each data point represents each model's ability to simulate both the placebo and intervention group's event rates.

The SimCare model had the most accurate prediction for the control group event rate (13.4% vs. 12.9%, CARDS and SimCare, respectively) and the second most accurate prediction for the intervention group (9.4% vs. 9.7%, CARDS and SimCare, respectively). SimCare also had the second closest prediction to the difference between the two rates of the intervention and control groups (4.0% vs. 3.2%, CARDS and SimCare respectively). The Sheffield model also generated similarly accurate

predictions, although little has been published about this model to grant insight into how or why these predictions occurred (see Appendix F).

ADVANCE – Blood Pressure

The ADVANCE-BP clinical trial was intended to show the effects of intensive blood pressure control on macrovascular and microvascular events (A. C. Group 2007). The effect of intensive blood pressure management was assessed using a routine combination of Perindopril and Indapamide or placebo in the two cohorts. Findings of the trial include that the drug combinations were well-tolerated and reduced the risks of major vascular events, including death. The SimCare study simulates the trial as 5 years of treatments to predict SBP values. Additional intermediate health outcome variables are also reported, as available, as are CVD event rates using the UKPDS Risk Engine.

Method

As was mentioned previously, the ADVANCE-BP trial was carried out in 20 countries, over several continents, and used 215 collaborating centers, making it a difficult trial to simulate accurately in every detail. Not only did patients span several ethnicities, treatment practices (e.g., evidence-based goals) and specific drug formularies were often left to local controls which varied. Table 20 shows the initial cohort characteristics in both the ADVANCE-BP trial and the SimCare simulated populations. Synthetic patients were randomly drawn from the original 10,000 patient super-population described in the introduction until the average health characteristics matched those in ADVANCE-BP.

Table 20. The initial parameter settings for each cohort in the actual and simulated populations for the blood glucose arm of the ADVANCE-BP trial simulation.

Initial Population Characteristics	ADVANCE-BP		SimCare	
	Active	Placebo	Active	Placebo
N	5569	5571	5569	5571
% Female	42.30%	42.60%	41.40%	43.90%
Age at Diagnosis (sd)	58(9)	58(9)	58(9)	58(9)
Duration of Disease(sd)	8.0(6.4)	7.9(6.3)	7.9(6.3)	8.0(6.3)
HgbA1c(sd) -- %	7.48(1.63)	7.48(1.65)	7.51(1.47)	7.52(1.55)
SBP(sd) -- mmHg	**145.0(21.4)**	**145.0(21.7)**	**145.4(19.5)**	**145.7(19.6)**
LDL(sd) -- mg/dL	120.4(39.5)	120.7(40.2)	120.4(34.8)	120.7(34.8)
HDL(sd) -- mg/dL	48.4(13.5)	48.8(13.5)	48.8(12.4)	48.8(12.8)
Median Triglyceride -- mg/dL*	145.1	141.6	159.3	157.5
Adherence	High	High	100%	100%

*Values of triglycerides in the simulated population had a higher median value due to ethnic disparities between ADVANCE populations and collaborating healthcare organization populations. For example, the median triglyceride values reported for SimCare (above) are generally similar to those reported in other studies conducted in the U.S. and Canada (Sacks, et al. 1996) (T. A. Group 2008), generally higher than those conducted in Asia and generally lower than those conducted in Scandinavian countries (H. P. Group 2003) (Gaede, Vedel, et al. 1999). Because no detailed information about the distribution was reported by the ADVANCE group, the original super-population distribution was used in the simulation.

Patients were treated by a simulated regimen that used evidence-based protocols (e.g., treat-to-goal) and defined drug dose formularies in order to simulate the treatment given to patients under the control of a clinical trial (Table 21). To match the scheduling intensities in ADVANCE-BP, patients in both cohorts made clinical visits two months apart for the first six months, and then one visit every six months after that. Because the ADVANCE-BP trial was conducted over such a wide set of countries and clinics, each with different practice standards and available medical formularies, a specific set of drugs was not published as the overall trial standard.

Table 21: ADVANCE-BP drug formulary available for blood glucose, blood pressure, and lipids panel management. Maximum doses are included in parentheses where appropriate.

Treatment Formulary	ADVANCE-BP		SimCare	
	Active	Placebo	Active	Placebo
Medication (max dose)				
Blood Glucose	Gliclazide		Glipizide (20mg)	
	Metformin		Metformin (2000mg)	
	Thiazolidinediones		Pioglitazone (45mg)	
	Basal Insulin		Basal Insulin	
	Prandial Insulin		Prandial Insulin	
Blood Pressure	Perindopril (8 mg)	Perindopril (0 mg)	Lisinopril (60 mg)	Lisinopril (0 mg)
	Indapamide (4 mg)	Indapamide (0 mg)	HDRZD(25 mg)	HDRZD(0 mg)
	Calcium Channel Blocker		Cardizem(240)	Cardizem(240)
	Beta Blocker		Atenolol(100)	Atenolol(100)
	ARB		Valsartan(320)	Valsartan(320)
Lipids Panel	Up to Local Physician*		Simvastatin (80mg)	
Scheduling				
First Six month	60 Days	60 Days	60 Days	60 Days
After 6 month	180 Days	180 Days	180 Days	180 Days
Goal				
A1c (Trigger)	Up to Local Physician*		7.0%(7.0%)	7.0%(7.0%)
LDL			135 mg/dl	135 mg/dl
SBP			140 mg/dl	145 mg/dl

*The ADVANCE-BP study did not publish these details. In order to simulate the trial, assumptions or substitutions for each area were made. For medications, collaborating physicians suggested viable substitute drugs within the same families of listed medications if the explicit drug used in the actual trial was not currently being modeled by SimCare. Evidence-based goals were derived empirically using trial and error.

The basic treatment algorithm was that if a patient was above the evidence-based goal for a given health characteristic, then an additional treatment move was administered. The goals given for this simulation were SBP <= 140 mmHg and 145 mmHg for the intervention and placebo cohorts, respectively. These evidence-based goals were derived empirically through trial and error by running multiple simulations using a range of clinically-acceptable goals and choosing each one that produced the mean outcome closest to the observed outcome. This method was used since no explicit goals were

published by the trial. The final simulation trial was then run using each of the goals discovered during the trial and error phase. Additional goals of LDL <= 135 mg/dL and A1c <= 7.0 % were set for both groups. Table 22 shows the intermediate health outcomes for all four cohorts after five years of treatment.

Table 22: Intermediate health outcomes published from the ADVANCE-BP clinical trial and simulated by the SimCare model. Results were tabulated after 5 years of treatment.

End of Follow-up	ADVANCE-BP		SimCare	
	Active	Placebo	Active	Placebo
N	5569	5571	5569	5571
SBP(s.d.)-mm Hg	**136***	**140**	**136.6(7.6)**	**141.0(10.9)**
HgbA1c (s.d.)-%	6.9	6.9	6.98(1.15)	7.15(1.26)
Tot Chol(s.d.)-mmol/L	5.0	5.0	4.50(1.06)	4.49(1.04)
HDL(s.d.)-mmol/L	1.00	1.00	1.28(0.34)	1.29(0.33)

*ADVANCE-BP did not publish standard deviations for these measures.

The main focus of the ADVANCE-BP simulation is to test SimCare's ability to generate cohort SBP levels accurately over time. The additional reporting of A1c and lipids levels is for the purpose of full disclosure (as it is also done in the published report for the trial). More detailed data for systolic blood pressure levels is reported in six-month increments for each cohort in Figure 12. Although some slight deviations between the simulation averages and the trial averages are seen within the first two years, the simulation converges with the trial for the last three years of comparison.

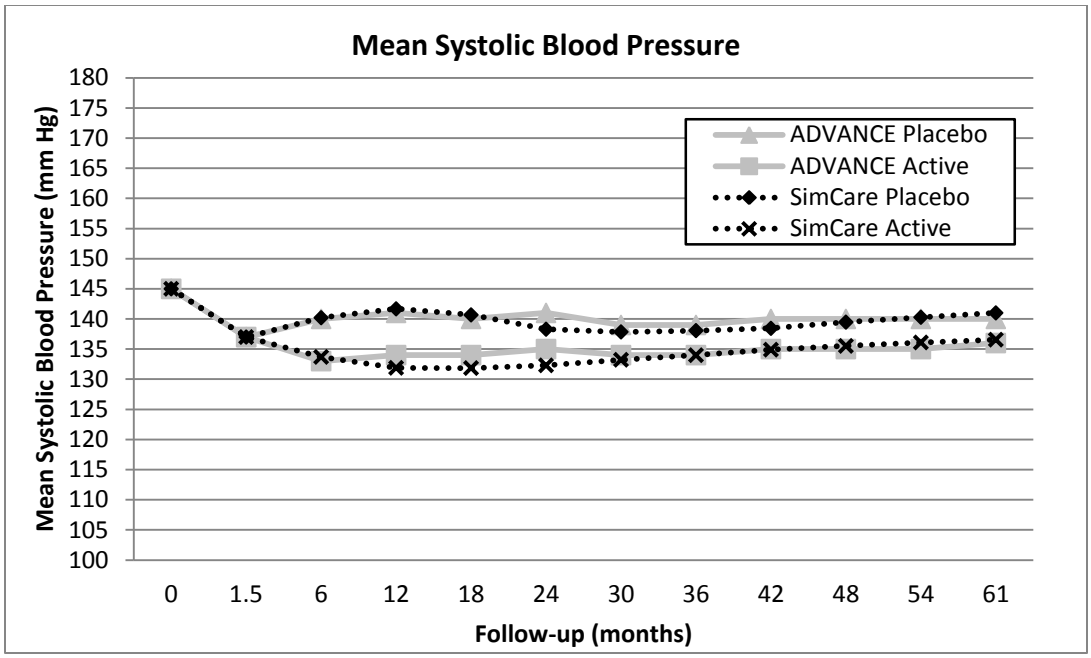


Figure 12: Mean SBP levels for each cohort (intensive and standard) in the ADVANCE-BP trial and SimCare simulation over the five years reported in the published trial.

An important detail published by ADVANCE-BP is the average SBP difference between the two cohorts. This difference was 5.6 mmHg over the course of the trial. The average SBP difference between the two cohorts simulated by the SimCare model was 5.59 mmHg. Additionally, a pre-randomization period in the trial was simulated by giving all patients low doses of indapamide (0.625 mg) and perindopril (2 mg) for the first six weeks of the trial. Upon randomization into two groups, patients were then given a double dose of the initial medications or the placebo. The effects of this pre-randomization detail can be seen in the figure as well.

Discussion

As mentioned in the discussion of the ADVANCE-BG simulation, the ADVANCE-BP trial represented both an important and difficult trial to simulate. The difficulty of

simulation stemmed from the heterogeneity inherent in a multi-country trial taking place in over 200 clinics, each with slightly different practices and available medical tools. Populations in different regions of the world with type 2 diabetes vary in certain baseline characteristics, especially triglycerides. This difficulty shows the importance of simulating the trial. Only basic simplifying assumptions were needed to run the simulation.

As seen in Table 23, the UKPDS Risk Engine (using SimCare data) over-estimated total event rates ranging from 0.4-1.1%. This consistent over-estimation might suggest that SimCare under-treated patients relative to ADVANCE-BP. However, this is not the case, as patient cohorts had nearly identical health outcomes with slightly over-treated lipids levels. These statistics would lead to the possible expectation of a lower estimation from the UKPDS Risk Engine (not higher, as seen). A possible explanation for this discrepancy is that ADVANCE-BP's undefined and broad category of "Macrovascular Events" fails to compare precisely with CHD and Stroke events (the only available event categories of output) from the UKPDS Risk Engine.

Table 23. Total event rates for the ADVBANCE-BP trial and predicted by the SimCare model using the UKPDS Risk Engine.

Primary End Points (rate)	ADVANCE-BP		SimCare	
	Active	Placebo	Active	Placebo
Macrovascular Risk*	480(8.6)	520(9.3)	542(9.7)	555(10.0)
Cardiovascular Death	211(3.8)	257(4.6)	268(4.8)	279(5.0)

*Macrovascular Risk was computed from UKPDS Risk Engine results by adding together CHD and Stroke events over the 5 year period.

Summary of Randomized Clinical Trial Validation

The SimCare model is a potentially powerful tool that can be used to simulate randomized clinical trials involving patients with type 2 diabetes. The use of SimCare to predict the results of randomized clinical trials tested the model's ability to predict three aspects of type 2 diabetes. These tests consisted of the ability of SimCare to: 1) predict intermediate health outcomes for defined cohorts, 2) predict differences in intermediate health outcomes between relevant cohorts, and 3) use a third party risk engine to generate various CVD total event rates.

Results

Cohort outcomes of treatment regimens of the SimCare Patient Model closely corresponded with observed outcomes of treatment for three main aspects of type 2 diabetes health markers—A1c, LDL, and SBP—for each relevant trial. Each comparison was created by taking the periodic, cohort mean data (usually every six months) from each cohort (i.e., group of patients in either arm of the trial) in the clinical trial that focused on that aspect of type 2 diabetes. For example, the CARDS trial was used for LDL comparisons and published 4 years of data for each year for each cohort (placebo and intervention) resulting in 10 data points for LDL comparison. The sample sizes for A1c, LDL, and SBP comparisons were 24, 10, and 24, respectively.

R^2 values are a way to calculate the error in a set of predictions relative to the underlying variance in observed outcomes and were calculated for each set of comparisons using the formula:

$$R^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \bar{x})^2}$$

In this formula, *x* represents an observed outcome, *y* represents a predicted outcome, and represents the mean of observed outcomes. This general method of calculating error is called the coefficient of determination and relates the sum of errors between the observed system and the set of predictions (often a linear model) to the underlying variation in the observed system. In a purely statistical context, an R^2 value measures the variation that a linear model – which was fit to an observation data set using regression – accounts for in the observation set. In this case it is related to the correlation coefficient and ranges between 0 and 1. In cases where the R^2 is computed from a general set of observations and predictions (e.g., from a non-linear model), the value is unbounded less than one. R^2 values equal to 1 indicate a set of perfect predictions. The closer an R^2 value is to zero, or the more negative it is, the greater the indication of inaccuracy in predictions. The R^2 values for each comparison are 0.9242, 0.9662, and 0.8887 for A1c, LDL and SBP, respectively. Additionally, HDL and Triglycerides were compared using the same method (for completeness) resulting in R^2 values of -2.2744 and 0.9070, respectively. Figures 13, 14, and 15 show the observed cohort means for A1c, LDL and SBP intermediate outcomes plotted against the predicted cohort means on top of the perfect prediction line. If a data point is above the line, that indicates the SimCare model "over-

predicted" the value; if the data point is below the line, the outcome was "under-predicted"; data points on the line represent a perfect prediction.

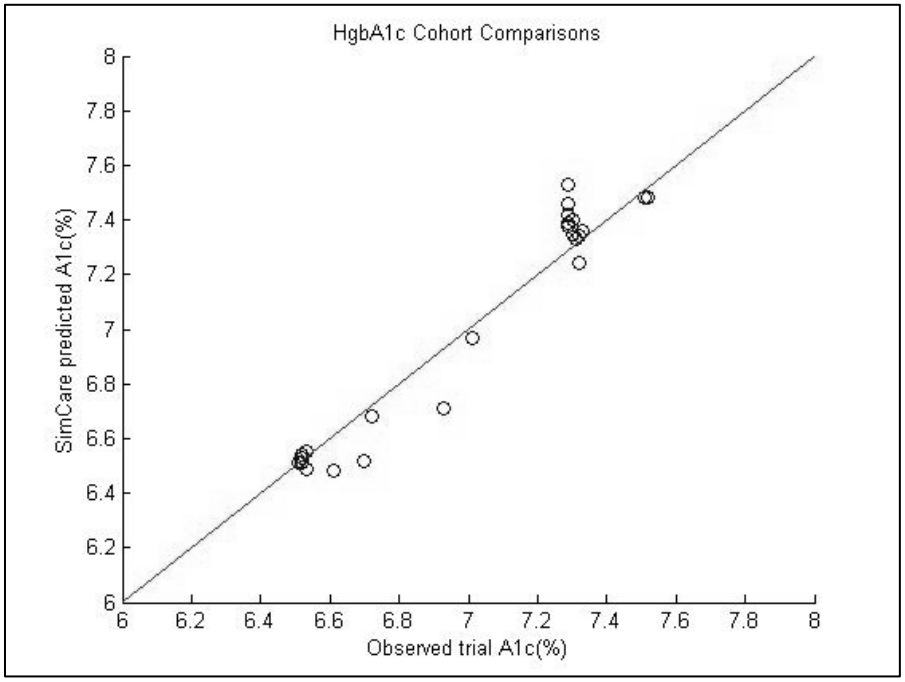


Figure 13. Scatter plot of observed and predicted average A1c levels for each cohort at each major time point (N = 24). The straight line indicates a perfect prediction by the SimCare model.

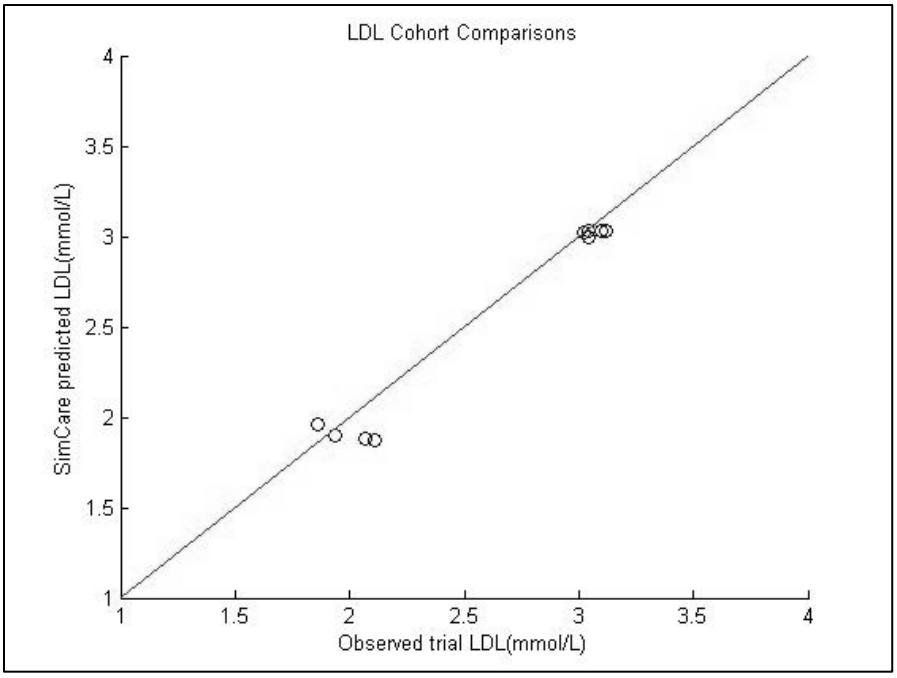


Figure 14. Scatter plot of observed and predicted average LDL levels for each cohort at each major time point (N = 10). The straight line indicates a perfect prediction by the SimCare model.

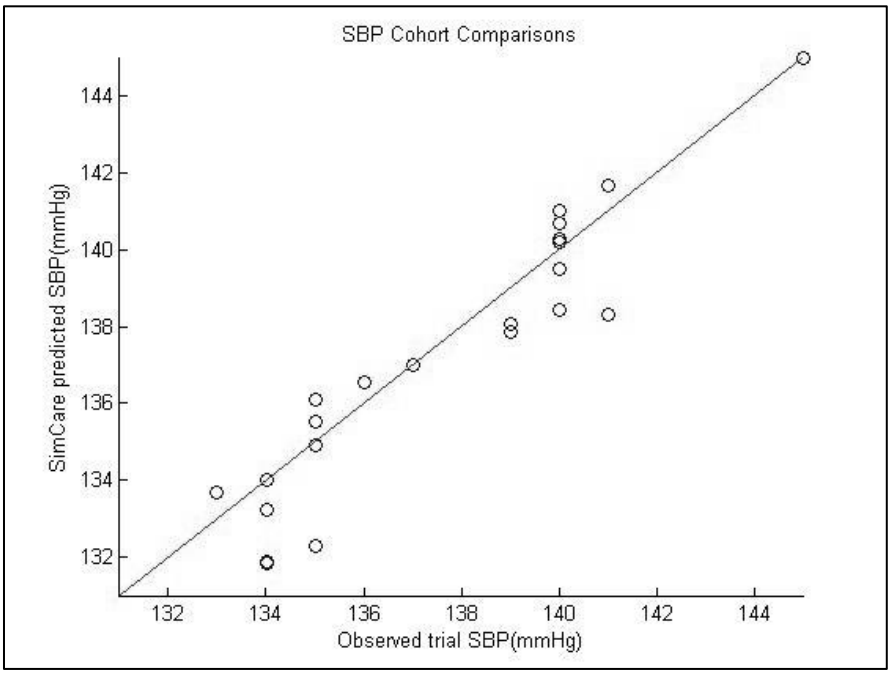


Figure 15. Scatter plot of observed and predicted average SBP levels for each cohort at each major time point (N = 24). The straight line indicates a perfect prediction by the SimCare model.

One of the main purposes of a randomized clinical trial is to demonstrate the differences between two cohorts of patients. In addition to accurately predicting cohorts separately, it is also important that an accurate prediction be made for the differences between cohorts. In order to form a basis for such comparisons across many different trials, a relative risk was computed between the control and intervention groups for the duration of each trial, using the detailed data from each trial. Relative risk is defined as a percent and is computed using the intermediate health outcome from the two comparative groups: (control – intervention)/control.

The calculation was done for each simulation study as well as for the data from each published trial. The comparisons of predicted and observed relative risk can be seen in Figure 16. An R^2 was computed for this data (N=44) with the resulting value = 0.8614. Four obvious outliers are clustered together near the lower left quadrant of the data. These four data points represent SimCare's predictions of HDL levels throughout the CARDS trial. As noted earlier, SimCare was unable to accurately predict these levels due to increased amounts of Atorvastatin and Simvastatin administered by the simulated, evidence-based protocol. If the HDL values are removed from consideration, the R^2 value rises to 0.9274.

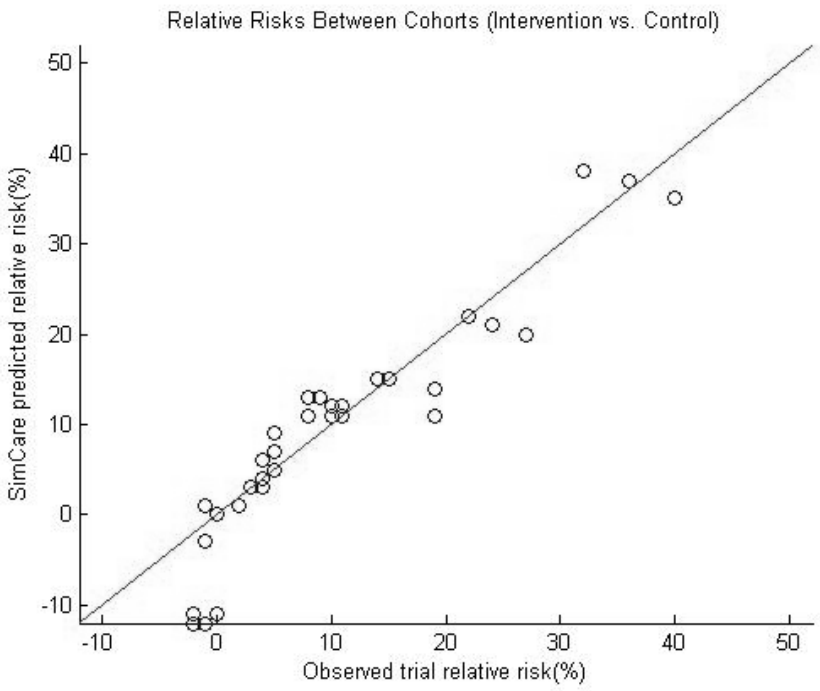


Figure 16. Scatter plot of observed and predicted relative risks between cohorts. Relative risk is defined for intermediate health outcomes as (Control – Intervention)/Control.

An additional concern of nearly every randomized clinical trial for type 2 diabetic patients is the effect treatments have on cardiovascular disease event rates. Event rates for various events (such as coronary heart disease or stroke) are often published in the form of total event rates = (total events)/(number of patients). Because the SimCare model does not generate CVD events, a third-party risk engine was used to generate risk profiles for each simulated patient in each trial. A Monte Carlo sampling process was used to generate specific events, and total event rates were computed. The ability of SimCare to generate intermediate health outcomes to inform a third-party risk engine so that simulated total event rates could be compared to actual total event rates is seen in Figure 17. The R^2 for this set of predictions was 0.9678 with N=12.

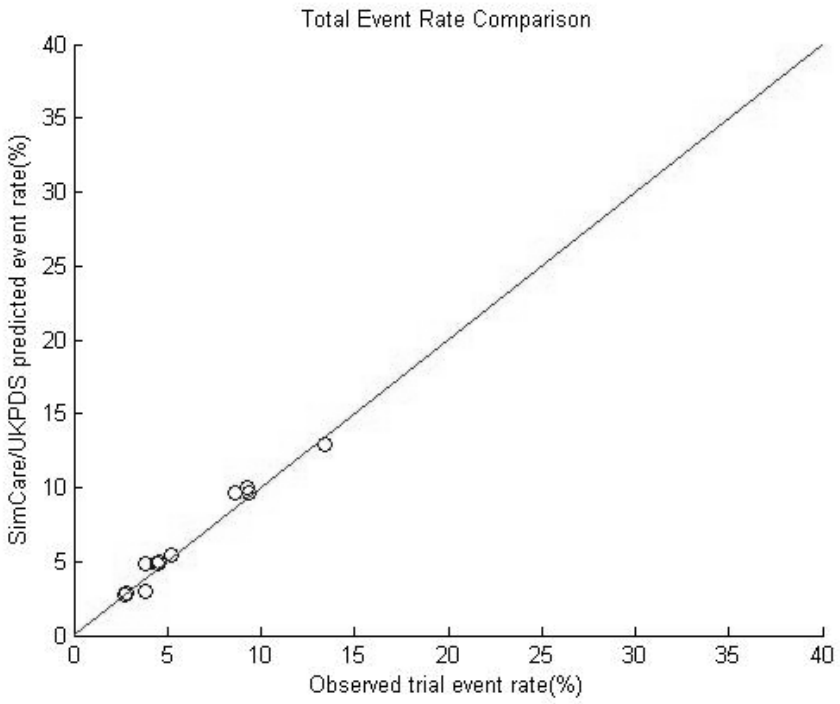


Figure 17. Scatter plot of observed and predicted total event rates. The UKPDS risk engine used SimCare data to simulate relevant events throughout each randomized clinical trial.

Sources of Error

Several possible sources of error exist in using the SimCare model. First, creating simulated populations to match those in published randomized clinical trials can be a difficult task for several reasons. Often, limited amounts of information for these patient cohorts is published, usually a mean and standard deviation for each relevant demographic and health factor. This could lead to the assumption that the population is normally distributed, which is true for many naturally occurring statistics (e.g., heights of men) but may not be true for the kinds of data represented in clinical trial populations (e.g., A1c levels in diabetic patients). Another difficulty in creating simulated

populations is that populations differ (e.g., triglycerides) due to ethnic and cultural differences. Even if ethnicity is not being modeled explicitly, the relevant distributions underlying disparate ethnic populations creates a need for additional descriptive population data. These relevant population descriptions are usually not addressed in detail by the published trials.

Part of simulating a randomized clinical trial—after creating a matching population of synthetic patients—is to establish a treatment regimen that accurately represents the treatments that occurred during the trial. There are several possible sources of error in this. One is accounting for differences in medical cultures and national drug approvals or usages. Not all drugs are available in all countries, so substitutions of modeled drugs within the same family of medications must be made with special attention given to matching dosing schedules. Also, medical cultures vary in medical practices from country to country as well as within a country; medical culture can vary from clinic to clinic within a single healthcare organization (Curoe, Kralewski and Kaissi 2003). These disparities remain as difficult to report in real trials as they are to simulate.

An additional source of error is possible when establishing treatment regimens for how physicians treated ancillary aspects of the complex patient. If the study was designed to test the effects of a lipid-lowering medication, but all the patients had type 2 diabetes, then how patients' blood sugars were treated throughout the trial must also be assumed and standardized for simulation. In a similar way, scheduling is often not precisely reported, showing a general range of clinical encounters suggested over the course of a

year with no detail of how many encounters occurred for patients outside of trial protocols. These encounters are often with the patient's primary care physician and could include treatments in addition to those in the trial protocols.

The SimCare model does not represent the occurrence of microvascular or macrovascular events for patients. These events were simulated using the UKPDS Risk Engine. The UKPDS Risk Engine was parameterized using intermediate health outcome data generated by SimCare trial simulations, but the risk engine has its own sources of errors and bias. For example, specific treatments for type 2 diabetes are not input, and only fatal and non-fatal, stroke and CHD events are predicted. The range of possible CVD events reported per trial is broader and often more detailed. Additionally, the entire category of microvascular events – of which A1c is believed to be an important determining factor (Lachin, et al. 2008) – is omitted from the UKPDS Risk Engine, thus limiting the scope of SimCare's predictions.

Given these possible sources of error, only one consistent bias emerged in the simulations, which was due to the way specific treatment protocols were simulated. As was mentioned for each simulation, an evidence-based protocol was used by establishing a treatment goal, a dosing schedule and a visit frequency. If a simulated patient had not yet reached the goal level at a given encounter, then the simulated protocol increased the relevant dose of the current drug or initiated an additional treatment if the previous drug had reached the maximum dose of clinical effectiveness. This simplistic protocol could result in a simulated patient receiving increased medication at a given encounter for each

of the three main intermediate health outcomes (blood glucose, blood pressure and lipids), which in practice is unlikely to occur. Additionally, a simulated patient with a very high level of some intermediate health outcome could receive monotonically increasing treatments at every single encounter, which in the definition of the trial protocols would also be unlikely. Either of these phenomena tended to decrease the amount of inter-patient variation over time in blood pressure and lipid variables due to the over-treatment of patients with high initial values. This effect showed up very clearly in the poor prediction of HDL levels in synthetic patients being treated with statins for high LDL. In order to improve upon this area, the simulated treatment protocol may need to separate the parameterization of each aspect of treating the patient model as well as incorporate some form of prioritization or time sensitivity.

Clinical Care Variation Validation of the SimCare Model

The SimCare model is a clinical model of an individual patient with type 2 diabetes. Instances of the model—representing individual patients—are treated on the occasion of clinical encounters over time. The results of these treatments are then stored in a database per patient, per encounter in a form similar to clinical medical records. The representation of individual patients with type 2 diabetes responding to treatments on a per encounter basis is one of the central claims of the computational theory reflected in the SimCare model and is the motivation for this N=1 study.

As stated previously, the operational validation of the SimCare model is carried out in two studies. The first study compared results of simulated randomized clinical trials to the published trial results in order to test the way the model represents aspects of cohorts of patients, including A1c, LDL, HDL, Triglyceride, and SBP. The second study uses clinical encounter data for individual patients and a canonical treatment model to test SimCare's ability to simulate variation over time for individuals with type 2 diabetes.

There are too many variables in the world of clinical care to explicitly simulate in each patient. Furthermore, even with a complete clinical data set, assumptions need to be made about basic conditions of treatment as well as unseen aspects of the state of the patient (e.g., adherence levels, biophysical response to medications, etc.). Since such underlying feature values cannot be observed in patient data – which would permit an experiment designed to simulate an exact treatment and outcome match per patient – they must be assumed as a range of values limited by model claims.

Another way to describe the context for this set of validation experiments is to see the problem from the perspective of a physician treating two patients who appear identical at a point in time in every clinical observation (e.g., A1c, LDL, etc.). If the physician gives both patients the same treatment regimen (e.g., 500 mg Metformin) and follows up with each after the same amount of time has expired, the physician would expect to see different effects from the treatment manifested in each patient at the follow up time, without necessarily being sure of the specific reason for the variation. Validating the variation the model is capable of generating (from a variety of causes) is the purpose of the N=1 set of simulation experiments. Explicitly, the experiments are designed to validate that the model does not over- or under-represent variation in a given patient (based on stated assumptions about that patient and experiment parameterizations) as well as expose any bias in particular areas of model domain space.

Each N=1 experiment reduces the scope of the comparison to the individual patient level. Because a set of assumptions must be made about the unobservable patient state for each patient, the model is parameterized under two sets of assumptions which serve as "book-ends" for a range of possible assumptions, in order to show variation boundaries for outcomes given each patient (see Figure 18). The first set assumes the best possible outcome given the highest adherence levels, smallest rates of disease progression, and the greatest response to treatments and lifestyle interventions. The other set of assumptions parameterizes the patient model such that the unobservable patient state variables reflect the worst possible clinical scenario. This assumption shows the outcome of the same

treatments on the same simulated patient and generates the worst possible outcome (i.e., the highest A1c %). The set of specific treatments for the patient found in the database is then used to duplicate a treatment protocol in the simulation experiment.

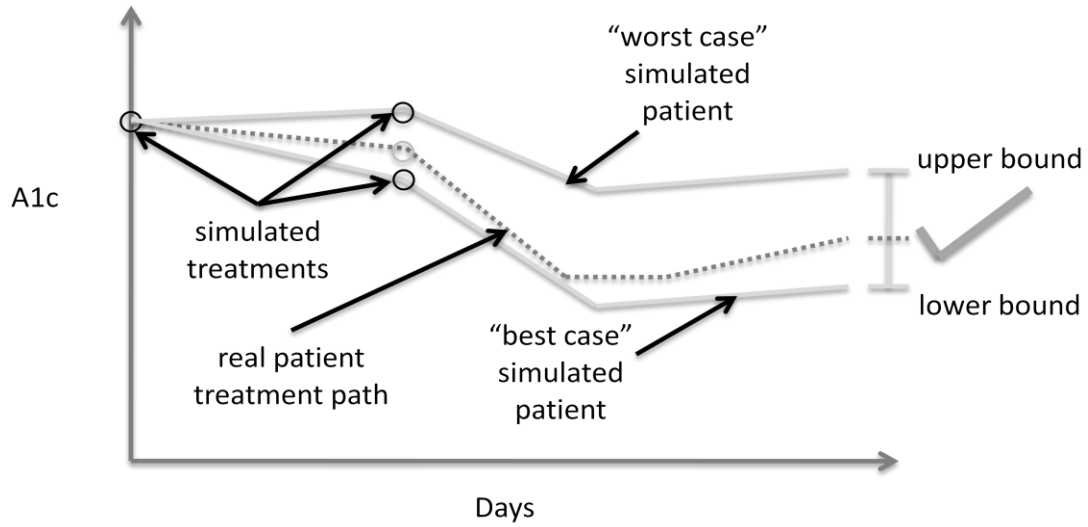


Figure 18. A conceptualization of the N=1 patient validation approach. Each real patient will be simulated twice as a best- and a worst-case scenario. The scenarios are established by setting the unobservable features of the patient state to the bounds of typical variation built into the SimCare model.

The intermediate and final outcomes (A1c values) from the experiment are recorded. The real patient outcomes are compared to these best and worst case outcomes to determine if they are within the simulated range of patient variation (i.e., between the best and worst case) to be a successful experiment. Patients outside of this simulated variation range are counted as experimental errors. Error rates are then totaled across the patient population and broken into sub-groups based on initial patient states and forms of treatment to determine if the model exhibits bias for a type of patient or treatment. The overall error rate is then used to test the hypothesis that the SimCare Patient Model can represent 95% of variation in individual patient outcomes.

Introduction

N=1 literature has been in use in the medical community for various reasons, mostly in an attempt to establish optimal treatments for individual patients (Guyatt, Sackett and Taylor, et al. 1986) (Mahon, et al. 1996). The idea within this literature is to conduct a personalized control and intervention series of treatments within a single patient and to use responses to those treatments in order to find the most suitable regimen for that patient (Guyatt, Sackett and Adachi, et al. 1988). The N=1 approach of this validation study is also centered on the idea of counterfactual thinking and learning from populations of one or fewer examples of events (Roese 1997) (March, Sproull and Tamuz 1991). This shows the value of learning from changes in the given contexts and assumptions of a problem that can affect outcomes. Beyond the modeling of known events is the simulation of unknown or unobserved events. The SimCare model is a tool for conducting such experiments. In order to have confidence in SimCare's ability to answer such questions, a demonstration of meaningful variation in individual patient outcomes over a range of real-world canonical patients (outside of controlled trials) and treatments must be established.

There are three aspects to this study. The first is to qualify the claims that SimCare models individual patient variation. This claim is limited to patients who go through canonical or standard sets of treatments and concomitant reactions. In the clinical world, certain patients may have a unique set of circumstances that would fall outside the claims of the SimCare model, for example, if that patient undergoes bariatric surgery or is

simultaneously battling cancer. These types of atypical patients may have highly variable reactions to treatments due to such exogenous reasons. These experiments are defined by a standard treatment algorithm with patient responses to cover only modeled cases of type 2 diabetes patients and treatments.

Secondly, there is the need to identify potential biases in the SimCare model. Such biases might manifest in simulated patients with very high initial A1c levels (e.g., A1c > 11%) or in patients receiving low doses of an oral medication, such as Metformin. Conducting the N=1 set of experiments over the patients spanning the entire space of the model's domain indicates where the model has a representative range of individual patient variation and where it does not.

Finally, this effort points out the usefulness of the various modeled aspects of individual patient variation in the SimCare model. The variation in question is not a function of randomness, which could be arbitrarily inserted in order to match any given set of outcomes, but is a function of observed variation in real patients and is due to real clinical treatments and outcomes. This meaningful variation enables the SimCare Patient Model to be used with confidence for both explanatory and predictive efforts such as presenting individual patient studies to physicians-in-training or optimizing treatment regimens for patients.

Clinical Care Data Set

A de-identified set of clinical encounter data spanning the dates March 1, 2005 to March 9, 2006 was used for purposes of the study. The previous year of data for this same data set was used to collect patient state variables (e.g., A1c) and medications such that patients were initialized with appropriate values for their first clinical encounter during the active time period. The resulting dataset included the patients who had been diagnosed with type 2 diabetes and included only those clinical encounters coded for diabetes treatments. The initial data set contained 2,174 patients over a total of 12,724 clinical encounters. None of these data had been used to develop the SimCare model.

The encounter-level data consisted of 11 variables necessary to conduct the experiment. Because the main focus of individual patient variation in the SimCare model is in managing A1c levels over time, lipid and blood pressure data were not included for this study. Each record was drawn from a larger patient database, and only records that had been flagged under the ICD-9 coding system as diabetes-related were used. The data had been de-identified with random identification numbers which were used to ensure the data from each encounter were assigned to the respective patient. In this way, both physicians and patients in the data were anonymously maintained. The age of the patient and encounter date were listed. The most recent A1c was listed. If a new A1c was not available for a given encounter, then the A1c value from the previous encounter for that patient was carried forward. Additionally, doses of oral medications Metformin, Sulfonylurea, and TZDs were listed per encounter. Insulin was also shown to be active

or inactive in patients, although exact doses were not available due to the way insulin is clinically administered to individual patients. The presence of insulin was indicated as a '0' or '1' entry for entries of Lispro, Lantus, and Other Insulins.

Canonical Forms of Treatment

The healthcare community offers several forms of protocol to guide physicians in the treatments of type 2 diabetes. A consensus algorithm from the literature developed by clinical practice physicians was used for the present N=1 study (Nathan, et al. 2009). The five forms of canonical treatments, shown in Figure 19, were used to categorize or exclude patients over the year of treatment observed in the database.

All patients were assumed to have been referred to a nurse educator – for Lifestyle Intervention (LI) such as diet and exercise education – since no data recorded that event, and it is a standard procedure upon diagnosis of the disease. The other categories were based on adding additional medications to the Lifestyle Intervention initiation. According to the treatment algorithm, patients who received LI might additionally be given Metformin if not at evidence-based goals. If the addition of Metformin was insufficient to bring the patient to the evidence-based goal, then a Sulfonylurea could be added. If that proved insufficient, then either a Thiazolidinedione or Insulin could be added. Medications need not have been fully titrated (i.e., reaching the maximum clinically effective dose) in order to justify starting a new medication. Multiple

medications could also be added simultaneously. In this way, canonical treatments are mutually exclusive (i.e., patients could not be in multiple categories at the same time).

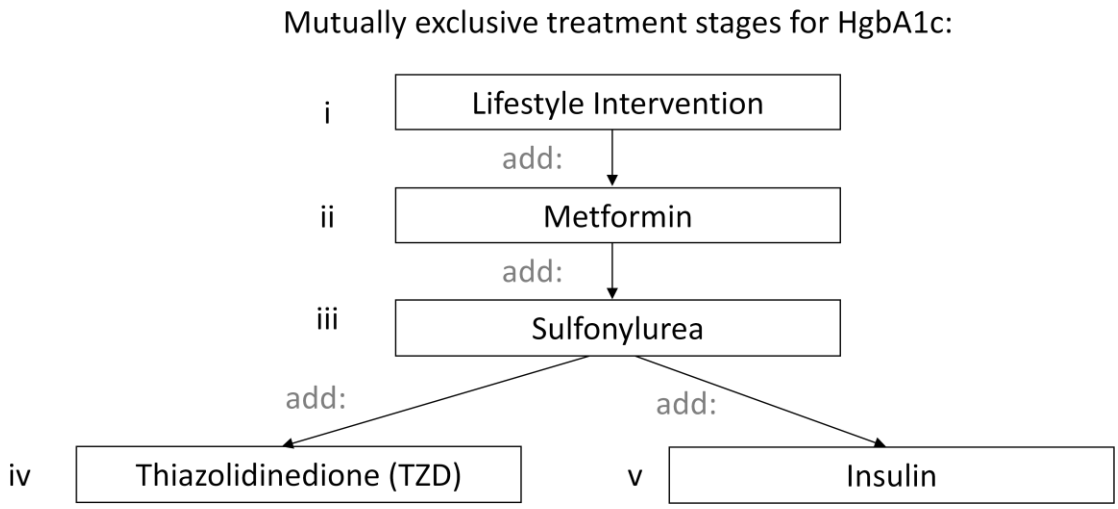


Figure 19. Five canonical forms of treatment used to categorize patients with type 2 diabetes. All patients included in the N=1 validation set started the year of treatment in one of these categories and finished in one of these categories. Each of the five forms of treatment is additive and mutually exclusive (e.g., a patient could start on Lifestyle Intervention only (i), and end up on Lifestyle Intervention + Metformin + Sulfonylurea + Insulin (v)) (Nathan, et al. 2009).

The first step in identifying patients to be simulated in the N=1 experiments was to categorize patients by the form of treatment with which they began or ended the year. Patients with non-canonical treatments (e.g., only on insulin) either at the beginning or end of the year were rejected from the N=1 experiment set. Intermediate patient encounter records throughout the treatment year were not considered in this process. A total of 219 out of 2,174 patients were rejected because their treatments were atypical.

Secondly, a set of boundaries was constructed to identify typical vs. atypical responses to canonical treatments. These boundaries were intentionally designed to include the greatest amounts of variation by adding together the maximal published effects of each

stage of treatment (Nathan, et al. 2009). In practice, individual physicians might see a smaller impact from combining several treatments. The bounds of each form of treatment are shown in Table 24. Values are negative for increasing treatments (i.e., A1c is expected to decrease) and are positive for decreasing treatments (i.e., removing medications from a patient would cause their A1c to rise). The full expected values are based on the full effects of the maximum clinically effective dose for each additional treatment stage after 90 days (the time it takes for oral medications to fully affect A1c levels). For example, 2000 mg of Metformin might lower a patient's A1c by as much as 2.0% after it has had the chance for the full effect to manifest in the patient's system (Garber, et al. 1997).

Table 24. Range of expected changes in A1c values compiled from the canonical treatment algorithm's published expectations (Nathan, et al. 2009). Bounds were generated conservatively to exclude patients exhibiting extreme variation over the course of treatment. Physicians in practice might expect to see less variation than listed in the table.

A1c Expectation Range (%)		End Phase				
		Lifestyle	LI + Metformin	LI + Met + Sulfonylurea	LI + Met + Sulf + TZD	LI + Met + Sulf + Insulin
Start Phase	Lifestyle	+1.0 to 0	-2.0	-4.0	-5.4	-7.5
	LI + Metformin	+2.0	+1.0 to 0	-2.0	-3.4	-5.5
	LI + Met + Sulfonylurea	+4.0	+2.0	+1.0 to 0	-1.4	-3.5
	LI + Met + Sulf + TZD	+5.4	+3.4	+1.4	+1.0 to 0	n/a
	LI + Met + Sulf + Insulin	+7.5	+5.5	+3.5	n/a	+1.0 to 0

Each patient corresponded to a cell in the table based on the canonical form of treatment they received as evidenced in the first and last clinical encounter for the year. A delta A1c for each patient was computed and compared to the limit in that cell. If a patient's

delta A1c exceeded the positive value or was lower than the negative value, then that patient was rejected from the experiment. Such large swings in patient variation could be explained (in reality) by many factors, such as exogenous diseases or treatments, unusual timing in A1c testing, or rare hyper-sensitivity to treatments. 895 out of the remaining 1955 patients were rejected for exhibiting extreme responses to canonical forms of treatment.

Finally, patients were considered based on the amount of information about their treatments that was recorded during the year. Many patients had three or fewer encounters in the dataset (e.g., a single encounter over the course of the year). Also, some patients showed evidence that a treatment move had been made, but then a follow-up result showing the effects of the treatment did not exist, making simulation for these patients difficult. 554 of the remaining 1060 patients were rejected from the N=1 simulation experiments for not having a sufficient amount of data.

In total 1668 of the original 2174 patients were rejected, leaving 506 patients for simulation. (See Summary section for discussion of rejection rate.) The occurrence rates per canonical treatment path of these patients are shown in Table 25. The majority of all patients (66.7%) began on no medications with only Lifestyle Intervention. A smaller number of patients (50.5%) ended the year (columns) on both Lifestyle Intervention and Metformin. The most populated cell (31.4%), indicating a treatment path over the course of the year, contained patients who began the year on Lifestyle Intervention alone and were then started on Metformin at some point during the year of treatment.

Table 25. The number of real patients to be simulated in each category of canonical treatment path. This table shows the final set of real patients to be simulated in the N=1 experiments. Rows represent how patients started the year; columns represent how patients ended the year.

Number of Patients in Each Canonical Group		End Phase					
		Lifestyle Intervention	LI + Metformin	LI + Met + Sulfonylurea	LI + Met + Sulf + TZD	LI + Met + Sulf + Insulin	
Start Phase	Lifestyle Intervention	75	159	73	19	13	339
	LI + Metformin	0	98	17	3	2	120
	LI + Met + Sulfonylurea	0	0	38	4	0	42
	LI + Met + Sulf + TZD	0	0	0	3	0	3
	LI + Met + Sulf + Insulin	0	0	0	0	2	2
		75	257	128	29	17	506

Generating Simulated Patients

After the patient cohort had been established, two sets of simulated patients were created. Each set of simulated patients matched every aspect of the real patient set that was established by the clinical database. Each simulated patient was assigned a number for identification, a set of encounters with identical dates to the real encounters, an initial A1c value and a set of treatments per encounter as listed in the real database. This ensured each simulated patient would be treated identically to each real patient.

In order to create the best- and worse-case scenarios used to conduct the experiments, the independent sources of variation in individual patient outcomes in the SimCare model were identified in the unobservable patient state set of features. Each variable was set to its highest or lowest value according to observed population confidence intervals. Although it is possible that higher or lower values existed in real world populations, they are unlikely and lie outside of the computational theory of the SimCare model, which is being tested by the hypothesis that the model can represent 95% of variation in individual patient outcomes. The variables were adherence, disease progression, metformin responsiveness, glipizide responsiveness, insulin responsiveness and lifestyle intervention (shown in Table 26). There are no known relationships among these patient features in nature, therefore the variables are modeled independently and parameterized individually.

Table 26. The parameterization settings used in the SimCare model to create the best- and worst-case scenarios for each patient simulated in the N=1 experiments. Means and descriptions of typical population distributions are listed to characterize the parameterizations.

Parameterization of N=1 Simulations	Best Case	Worst Case	Mean	Distribution	Source
Adherence (%)	99	30	87	Exponential	(Organization n.d.)
Disease Progression (A1c %)	0.004 (year)	1.61 (year)	0.35 (year)	Mixture	(Levy, et al. 2004)
Metformin Responsiveness (%)	112	88	100	Normal	(Garber, et al. 1997)
Glipizide Responsiveness (%)	127	73	100	Normal	(Campbell 1988)
Insulin Responsiveness (%)	145	78	100	Normal	(Esposito, et al. 2008)
Lifestyle Intervention	Effective	Not Effective	n/a	Boolean	(Rickheim, et al. 2002)

Adherence models the concept of how well a patient adheres to the medical regimen, dietary advice, and any other form of treatment established between the doctor and

patient. Adherence is a psychosocial variable affected by a range of factors from lack of belief in the importance of treatment to the ability to afford medication. Disease progression is a variable that describes the natural, physical progression of the disease if left untreated. The responsiveness variables are set separately for disparate families of medications and represent the natural physiological variation across patients in response to identical treatments. Dose response curves in the SimCare model are defined and fixed per dose from medical literature. Any variation of responses (above or below the mean response) across patients due to physical differences is modeled in the responsiveness variables. For example, as is shown in Table 26, 95% of patients respond to Metformin within an 88-112% response of the mean response. Lifestyle intervention is an additional variable that affects some patients and others not at all. Because it is not known for each patient when and if Lifestyle Intervention was used by the treating physician, it was either assumed to have a full effect or no effect for each scenario.

As shown in Table 26, the best- and worst-case values used to create the simulated patients were derived from the confidence intervals provided in published sources. For example, the responsiveness values were determined by deconstructing the confidence intervals around published dose response curves. Dose response curves are reported as average responses, which is why the average responsiveness is 100%. Individual patients may, however, receive a smaller or larger benefit from the respective drugs depending upon their level of responsiveness to each drug. Characterizations of the distributions were disclosed for information's sake. The columns of values of best- and worst-case scenarios were used to parameterize, respectively, each set of simulated patients in the

best- and worst-case simulations under the experimental assumption that 95% of variation in patient outcome could be captured by the model. Once a simulated patient's values were initiated, they did not change throughout the course of the experiment.

An example of a typical N=1 experiment is shown below in Table 27. The patient had an initial A1c = 8.6% and eight clinical encounters over the course of the year. This patient started the year in the canonical treatment form of Lifestyle Intervention only and ended the year with LI + Metformin + Sulfonylurea (Glipizide). Encounter dates are expressed in terms of days from the initial encounter, which are started at day 0 for each patient. A1c values for each patient were carried forward from previous encounters or entered in bold when new information became available (i.e., an A1c test had been ordered). SimCare can generate a current A1c value for any given encounter, thus all simulated A1c values are reported and are current for the patient at that encounter. This delay in real A1c information when compared to simulated A1c values at each encounter may create the perception of inconsistencies or errors in simulation from one encounter to the next. For this reason, only the final encounter was used to compare the real patient result with the simulated best and worst scenarios.

Table 27. An example patient from an N=1 experiment. The table shows the patient treatment path, real and simulated, for each clinical encounter.

Encounter Num	1	2	3	4	5	6	7	8
Encounter Date	0	21	105	166	182	245	273	282
Metformin (mg)	0	0	0	0	500	500	500	500
Glipizide (mg)	0	0	10	10	10	10	10	10
Pioglitazone (mg)	0	0	0	0	0	0	0	0
Insulin (0 or 1)	0	0	0	0	0	0	0	0
SimCare Worst Case A1c (%)	8.6	8.7	9.1	9.0	9.1	9.1	9.2	9.3
Reported Clinical A1c (%)	**8.6**	8.6	8.6	**8.4**	**7.1**	**6.5**	6.5	6.5
SimCare Best Case A1c (%)	8.6	8.4	8.3	6.6	6.6	5.5	5.5	5.5

The medication doses are listed in mg, with the absence or presence of insulin listed as a 0 or 1, respectively. Current A1c values for the best- and worst-case scenarios are shown as computed by the SimCare model on each encounter. Actual patient A1c values reported in the clinic are shown in bold when new information was available. The example in Table 27 indicates a successful simulation experiment as the real patient's final A1c value was between the upper and lower A1c bounds established by SimCare.

Results of Clinical Care Variation Validation

The initial patient space was divided into low (A1c <= 8.0%), medium (8.0% < A1c <= 10.0%) and high (A1c > 10%). The number of patients in each cell can be seen in Table 28. A trend can be seen in the simulation accuracy from patients in low, medium, and high cells. This trend could be expected as patients with higher A1c levels tend to exhibit greater amounts of variation as more medications are given to patients in the high cell; these patients are also more likely to be given insulin. The specific amounts of insulin

given to specific patients is not available in the patient record, so a conservative algorithm was used to administer insulin to simulated patients when indicated. Further evidence of this trend is seen in the average error in the small number of trial errors in the high cell, where A1c levels simulated for each best case over-estimated final A1c by nearly a full point (0.97%).

Table 28. The N=1 patient population broken into sub-groups based on initial patient states.

N=1 Simulations by Initial A1c	Count	Avg. Init HP A1c	Avg. Final HP A1c	Avg. Low SimCare Range	Avg. High SimCare Range	Patients Within Range	Percent	Patients Below Range	Avg. Error	Patients Above Range	Avg. Error
A1c <= 8	417	6.81	6.62	5.29	7.9	412	98.8%	1	-0.59	4	0.26
8 < A1c <= 10	71	8.88	7.98	6.86	9.91	66	93.0%	4	-0.30	1	0.07
A1c > 10	18	11.07	10.08	9.25	12.1	15	83.3%	3	-0.97	0	n/a
	506					493	97.4%				

Additionally, patients were broken into subgroups based on canonical forms of treatment. This subgroup analysis was used to determine if any bias existed for particular types of treatments. Table 29 shows the number of simulation trials in each treatment path that resulted in error. Although the model simulated variation for each group, an expected trend can be seen in the cells where the most medication was given over the course of the year (this is related to a patient's initial A1c as seen in the previous table). The more medication introduced to a patient's system in a fixed period of time (e.g., 1 year), the more opportunities exist for error to be introduced into the simulation. This is a straightforward effect of having more sources of variation in the patient model activated and combined.

Table 29. The N=1 patient population broken into sub-groups based on canonical forms of treatment. Rows represent the form of treatment on which a patient began the year. Columns represent the form of treatment on which a patient ended the year. Cells represent the treatment path for each patient throughout the year and contain the number of patients below and above the simulated range (in error) in addition to the percent of successful N=1 patient experiments.

Num Patients Below/Above Range (% Successful)		End Phase					
		Lifestyle Intervention	LI + Metformin	LI + Met + Sulfonylurea	LI + Met + Sulf. + TZD	LI + Met + Sulf. + Insulin	
Start Phase	Lifestyle Intervention	0/0 (100.0%)	3/0 (98.1%)	2/0 (97.3%)	1/1 (89.5%)	1/0 (92.3%)	7/1 (97.6%)
	LI + Metformin	n/a	0/2 (98.0%)	1/0 (94.1%)	0/0 (100%)	0/0 (100%)	1/2 (97.5%)
	LI + Met + Sulfonylurea	n/a	n/a	0/2 (94.7%)	0/0 (100%)	n/a	0/2 (95.2%)
	LI + Met + Sulf + TZD	n/a	n/a	n/a	0/0 (100%)	n/a	0/0 (100%)
	LI + Met + Sulf + Insulin	n/a	n/a	n/a	n/a	0/0 (100%)	0/0 (100%)
		0/0 (100%)	3/2 (98.1%)	3/2 (96.1%)	1/1 (93.1%)	1/0 (94.1%)	8/5 (97.4%)

The overall success rate for the N=1 set of simulations was 97.4%. The hypothesis of the experiment was that 95% of variation in patient outcome could be captured by the SimCare model. Testing the experimental error rate to the assumed error rate is a way to further validate the N=1 validation method used. If 100% accuracy had occurred in the N=1 study, it could be argued that the experiment had parameterized the SimCare model to include too much variation. Or, if the error rate had been too high, the overall N=1 study could be interpreted to show that the SimCare model does not represent the claimed amount of patient variation. There is some acceptable range of values for the error rate, given the set of experiments was based on 506 trials. An experimental error rate that is above or below this range would call into question the experimental design as well as the integrity of the SimCare model's ability to represent individual patient variation.

To bridge the gap between the novel experimental design and experimental result, a further, statistical simulation experiment was conducted. First, the *a priori* simulation assumption of capturing 95% of patient variation was hypothesized for the model and served as the null-hypothesis to test the experimental error rate. Then, a confidence interval around the assumed mean was established as 99%. If the experimental error rate falls outside of the confidence interval around the assumed error rate, then the null hypothesis would be rejected.

To generate an empirical distribution of error rates for a set of 506 experiments, a random number generator (using a uniform distribution) was used to simulate the outcomes (i.e., success or error) of 506 experimental trials using a 95% accuracy threshold. For each set of experiments, 506 numbers between 0-1 were generated and compared to the threshold of 0.95. If a number was over the threshold, then it was counted as an error, otherwise it was counted as a success. The error rate for each set of experiments was computed by summing the errors for the set and dividing by the number of trials (506). 5000 sets of these simulated experiments were conducted to empirically derive a distribution of error rates. This sampled distribution was characterized by a mean (0.9502) and a standard deviation (0.0095). A random simulation with a different number of trials or a different assumed distribution of randomness (e.g., non-uniform) would result in a different variance of expected experimental outcomes. In any case the central limit theorem predicts a normal distribution of error rates will be generated.

Expanding the mean by three standard deviation in each direction, a 99% confidence interval was created (0.9216—0.9788). The observed success rate of the SimCare N=1 trial (0.9743) is within this range (Figure 20) resulting in the conclusion that the null hypothesis cannot be rejected—that the observed result of the N=1 study could have come from an *a priori* assumption of predicting 95% of patient variation.

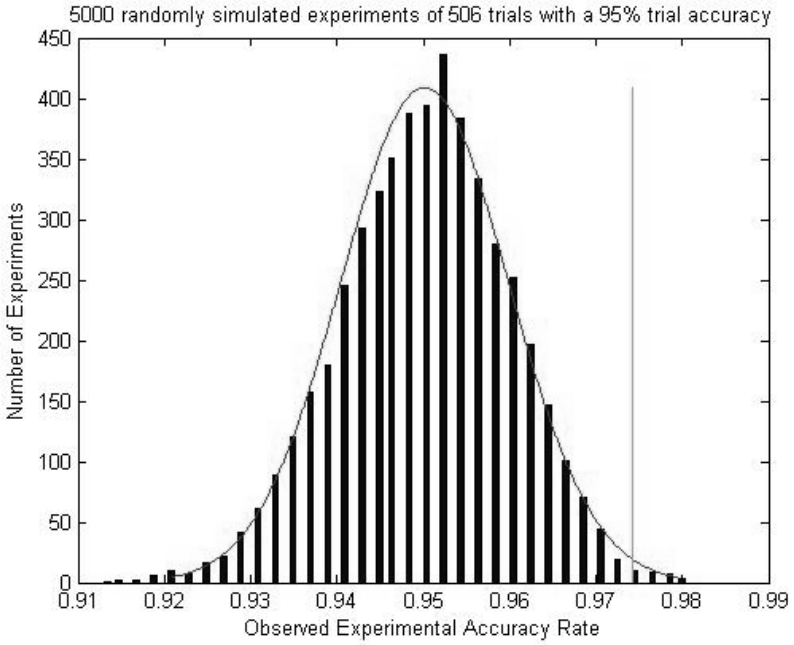


Figure 20. Histogram of a distribution of randomly generated success rates from an *a priori* rate of 0.95. The curve is the plotted normal distribution parameterized (and scaled) from the mean and s.d. of the data that generated the histogram and spans three standard deviations in each direction from the mean to span a 99% confidence interval. The vertical line is the observed SimCare success rate for the N=1 study.

The entire data set for simulation errors is shown in Appendix E. Overall, the A1c values of 13 patients fell outside of their simulated bounds. Five had A1c's higher than the simulated worst case; eight had A1c's lower than the best case. By analyzing these

thirteen cases alone, no clear trend or pattern could be seen in the types of patient in error. Patients whose A1c's exceeded the worst case tended to miss by a smaller margin of error (+0.22%) than did patients who fell below the best case (-0.59%). This disparity is not caused by the SimCare model directly but may indicate a bias of the parameterization used to set the best and worst cases. Such a bias could be the result of an underlying skew in distributions of published variables that were assumed to be normal.

Summary of Clinical Care Variation Validation

Using real clinical encounter data from 506 patients, SimCare simulated specific treatments on two sets of synthetic patients whose initial health states identically matched the initial health states of the real population of patients. These two synthetic sets varied in how their unobservable patient state features were parameterized in order to represent an upper and lower range of typical individual patient variation. Both sets of treatment paths and patient outcomes were constrained to represent only variation explainable by the SimCare model by creating a best and worst case scenario for each real patient. By the end of each patient simulation, the best and worst case scenarios had diverged and were compared to the observed real patient outcome. If the real patient's A1c level was between these two bounds, then the simulation trial for that patient was successful; otherwise, an error was noted.

This patient population was then divided into *a priori* subgroups along two relevant dimensions—initial patient state and canonical treatment path—to analyze the overall performance of the SimCare model and, specifically, to address the accuracy of its predictions throughout the simulation space. This two-fold analysis of results showed at least one trend that was related to limitations of experimental data (i.e., insulin doses were not recorded). Other trends that would indicate a bias in the SimCare model either in successful experiments or experimental errors were not found. The error rate for the set of experiments supported the hypothesis that the SimCare model can represent 95% of variation in individual patients.

The rejection of 76.7% of the initial raw data can be explained by several aspects of the experimental design and should not in and of itself cast skepticism upon the model or the validation study. This study used data from a primary care clinical healthcare practice which allows for types of error in data entry and informatics design that may not be as common in more rigorous healthcare research (e.g., randomized clinical trials). Also, a particular treatment algorithm was used to define the canonical forms of treatment in this experiment, and the clinical data set used was descriptive of a local population of people. Future validation studies could make use of data sets that go further in time as well as employ additional treatment algorithms.

Discussion and Conclusion

The work described by this thesis contributes to the understanding of computational modeling and validation by presenting a two-part validation of a model of patient illness (the SimCare Patient Model), a conceptual theory for a model of patient illness, and a form of N=1 operational validation for models of individuals with variation in outcomes over time. The first part of the validation was a conceptual validation; the second, an operational validation. The validation was conducted in two parts to satisfy two overlapping but different sets of concerns. The first set of concerns is in the set of representations the model has claimed. The second set of concerns is in the accuracy of the predictions the model generates.

The conceptual validation identified the research problem that was stated, namely, the sequence of changes of the health state of a patient with type 2 diabetes to the progression of the disease and a variety of outpatient administration of treatments by physicians over time. This validation provided a level of transparency to the workings of the model, so that its medical underpinnings are clearly defined in relation to their algorithmic representations. Because of the intent of the SimCare model, these representations are not opaque, mathematical functions, but straightforward, mechanistic definitions from medical literature. This mechanistic property relates what is in the model and why it is relevant to how it is represented and enables the model to provide relevant explanations to observed emergent behavior.

Two main constructs anchor the "what" and "why" conceptual theories of the model: the way type 2 diabetes is represented in the patient and the way clinical encounters are represented. These two constructs can be generalized across healthcare models of individuals with disease. The construct of disease representation – including observable, unobservable, direct and indirect feature sets – would be useful to define any disease state in an individual but would be particularly useful for complex patient states (diseases with multiple, separate but related treatable attributes) and with multiple sources and manifestations of variation in patient state changes. The construct of the clinical encounter would apply to a model that identifies the constraint of specific opportunities to treat individuals (e.g., case-based) and may include multiple opportunities for treatment depending upon changes in the patient state. Together, these constructs could guide the design and validation of conceptual models of chronic diseases very well, however they might also be appropriate for other patient illnesses that are currently not considered chronic diseases (e.g., patients with newly diagnosed or recurrent cancer).

For the SimCare model, which represents patients with type 2 diabetes, the constructs provide clear explanations to the healthcare community of problem definition, model content and model domain. The representation of type 2 diabetes is defined by a set of patient state health variables and the actions that can affect the patient state. The clinical encounter is defined to frame the context in which the patient state generates values to inform a decision-maker, registers any actions as a response to that state and computes the next patient state based on the timing of the next scheduled encounter. How these theories are represented is defined in the dynamics of the model pathways. These

pathways connect a given patient state to the generation of the next state through straightforward functional representations of dose and time response curves. These pathways define the way both the individual characteristics of the patient and the exogenous effects of treatment actions affect the computation of subsequent patient states. The conceptual validation of the SimCare Patient Model provides a level of explanation to enable two contexts for intended use: The way the model represents responses to treatments in controlled clinical trials, and the way the model represents variation in individual patients in a clinical care setting over time.

The operational validation of the SimCare model was designed around the contexts of its specific, intended uses. These uses were described in the conceptual validation chapter of this thesis, which provided a basis for the set of experiment simulations that enabled the predictions of the model to be compared to real observations. Two operational validation studies were conducted using simulation experiments to validate the model within these two contexts.

The first study was used to compare SimCare's ability to simulate randomized clinical trial outcomes to published clinical trial outcomes. Three clinical trials were selected for simulation to test three main aspects of the model – A1c, Lipids, and SBP. Synthetic populations of individual patients were generated to match statistical descriptions of each cohort on a trial by trial basis, as were drug formularies, encounter scheduling frequencies and evidence-based goals. These treatment regimens were then applied to each patient in each cohort and descriptive statistics were generated for comparison to

trial data. A third party cardiovascular disease risk generator – the UKPS Risk Engine – used simulated patient intermediate health outcomes as inputs and generated events throughout each trial simulation. The resulting simulated cohort data – cohort means, differences between cohort means (relative risks) and CVD event rates – were compared to published data so prediction errors could be used to characterize the goodness of fit of model predictions. Although several sources of potential error were identified and discussed, the only consistent bias in the model was not from patient state responses *per se* but from the parameterization of the way treatment regimens were simulated. If more detailed data were published by trial results, perhaps more accurate treatment regimens could be defined. This, as well as the use of additional risk models, would improve the ability of the SimCare model to predict the results of randomized clinical trials.

The second validation study was conducted by simulating a series of N=1 experiments. This approach to model validation is a novel approach constructed via this thesis for the purpose of validating the SimCare model in the domain of clinical care of individual patients. This approach can be generalized to validate models of patient illness where typical care in a clinical (or other) setting may vary from patient to patient for a number of reasons compared to general treatments applied to populations (or rote application of medical guidelines to individuals).

Additionally, this approach is useful in modeling problems where identical treatments can result in different outcomes for different patients with identical, observable disease states at some point in time. Because there are multiple, unobservable sources of variation for a

patient with complex disease in such models, there would be multiple ways of parameterizing a model to fit a given patient instance in hindsight (e.g., using a machine learning algorithm). This more narrow form of validation may show more about a given patient instance but would indicate less about the general capabilities of the model. Instead, by establishing *a priori* a set of typical patients and treatments and by establishing two sets of extreme parameters across the model variables that are responsible for variation in individual outcomes, a general claim can be validated for the model based on the way actual experiments might be run with available data and necessary assumptions (i.e., not knowing explicit values for every source of variation in a given patient). By making a general claim about variation in individual patients, such an N=1 validation is able to show a model doesn't under- or over-represent variation in patients based on given model parameters. Additionally, such a claim using an N=1 approach can validate the model across clinically meaningful patient categories – including patient states (e.g., high, medium, low A1c) and treatment types (e.g., drug families) – by identifying any particular bias throughout the domain space.

For the N=1 validation of the SimCare patient model, each experiment tested the ability of the model to represent variation in individual patient outcomes to typical treatments over time. The set of experiments was defined by a real clinical database that was used to establish for each patient an initial patient state, a specific timing of encounters, precise medication doses for each encounter and a final patient state for comparison. An initial assumption was made and the model was parameterized for each patient simulation to represent 95% of typical patient variation. This type of assumption is a necessary part of

conducting a set of N=1 experiments designed as such, for two reasons: 1) the SimCare model does not claim to represent all variation in all patients, and 2) ultimately, the SimCare model is a computational model used to conduct simulation experiments, and it must be shown that it can be parameterized *a priori* to achieve a predictable result. For each N=1 experiment, patients whose A1c values were within the best and worst case range at the end of each simulation were counted as successful experiments, otherwise as errors. The resulting error rate was compared to the assumed error rate to further validate the design of the set of experiments as well as its outcome. The resulting conclusion provided evidence that the SimCare model can represent 95% of clinically meaningful variation in individual patient outcomes over time.

Bibliography

Aronoff, S, et al. "Pioglitazone Hydrochloride Monotherapy Improves Glycemic Control in the Treatment of Patients with Type 2 Diabetes." *Diabetes Care* 23 (2000): 1605-1611.

Association, American Diabetes. "Standards of Medical Care in Diabetes -- 2011." *Diabetes Care*, 2011: S11-S61.

Bachorik, PS, and JW Ross. "National Cholesterol Education Program Recommendations for Measurement of Low-Density Lipoprotein Cholesterol: Executive Summary." *Clinical Chemistry* 41, no. 10 (1995): 1414-1420.

Bagust, A, P K Hopkinson, W Maier, and C J Currie. "An economic model of the long-term healthcare burden of type II diabetes." *Diabetologia*, 2001: 2140-2155.

Bechtel, William, and Adele Abrahamsen. "Explanation: a mechanist alternative." *Stud. Hist. Phil. Biol. & Biomed. Sci.*, 2005: 421-441.

Bishop, Christopher M. *Neural Networks for Pattern Recognition.* New York: Oxford University Press, 1995.

Boland, Elizabeth, Teresa Monsod, Maria Delucia, Cynthia A Brandt, Sanjay Fernando, and William V Tamborlane. "Limitations of Conventional Methods of Self-Monitoring of Blood Glucose: Lessons learned from 3 days of continuous glucose sensing in pediatric patients with type 1 diabetes." *Diabetes Care*, 2001: 1858-1862.

Brandle, M, and W H Herman. "The CORE Diabetes Model." *Curr Med Research and Opin (Suppl 1)*, 2004: S1-S3.

Brehmer, B. *Strategies in Real-Time Dynamic Decision Making.* Edited by R Hogarth. Vol. Insights in Decision Making: A tribute to Hillel J. Einhorn. Chicago: University of Chicago Press, 1990.

Campbell, R K. "Glimepiride: Role of a New Sulphonylurea in the Treatment of Type 2 Diabetes Mellitus." *Ann Pharmacother* 32 (1988): 1044-1052.

Carter, B L, M E Ernst, and J D Cohen. "Hydrochlorothiazide Versus Chlorthalidone Evidence Supporting Their Interchangeability." *Hypertension* 43 (2004): 4-9.

Ciechanowski, Paul S, Wayne J Katon, Joan E Russo, and Edward A Walker. "The Patient-Provider Relationship: Attachment Thoery and Adherence to Treatment in Diabetes." *American Journal of Psychiatry*, 2001: 29-35.

Clarke, P M, et al. "A model to estimate the lifetime health outcomes of patients ith Type 2 diabetes: the United Kingdom Propsective Diabetes Study (UKPDS) Outcomes Model (UKPDS 68)." *Diabetologia*, 2004: 1747-1759.

Colhoun, H, et al. "Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomized placebo-controlled trial." *The Lancet* 364 (2004).

Conant, R, and W R Ashby. "Every Good Regulator of a System Must Be a Model of that System." *Int. J. Systems Sci.* 1, no. 2 (1970): 89-97.

Cosmides, Leda, and John Tooby. "The Cognitive Neurosciences." Edited by Michael S Gazzaniga, 1199-1210. Cambridge, MA: The M.I.T. Press, 1995.

Cowie, Catherine C, and Maureen I Harris. "Diabetes in America." By Ronald Aubert, 117-164. National Institutes of Health, 1995.

Curoe, Ann, John Kralewski, and Amer Kaissi. "Assessing the Cultures of Medical Group Practices." *The Journal of the American Board of Family Practice*, 2003: 394-398.

Dahlof, Bjorn, et al. "Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol." *Lancet* (Lancet) 359 (2002): 995-1003.

De Jong, Hidde. "Modeling and Simulation of Genetic Regulatory Systems: A Literature Review." *Journal of Computational Biology*, 2002: 67-103.

DeFronzo, R A. "Pharmacologic Therapy for Type 2 Diabetes Mellitus." *Ann Intern Med* 131 (1999): 281-303.

Derr, R, E Garrett, G A Stacy, and C D Saudek. "Is HbA1c Affected by Glycemic Instability?" *Diabetes Care* 26 (2003): 2728-2733.

Dodson, P M, M Beevers, R Hallworth, M J Webberley, R F Fletcher, and K G Taylor. "Sodium restriction and blood pressure in hypertensive type 2 diabetics: randomised blind controlled and crossover studies of moderate sodium restriction and sodium supplementation." *Br Med J* 298 (1989): 227-30.

Donnan, P T, T M MacDonald, and A D Morris. "Adherence to prescribed oral hypoglycaemic medication in a population of patients with Type 2 dabetes: a retrospective cohort study." *Diabetic Medicine*, 2002: 279-284.

Dunbar-Jacob, J, and M.K. Mortimer-Stephens. "Treatment adherence in chronic disease." *Journal of Clinical Epidemiology*, 2001: S57-S60.

Dutta, Pradyumna, et al. *SimCare: A Model for Studying Physician Decision Making Activity.* Vols. Advances in Patient Safety: from research to implementation: programs, tools, and products, by Agency for Healthcare Research and Quality. Rockville (MD): AHRQ, 2005.

Earnshaw, S R, et al. "Optimal Allocation of Resources across Four Interventions for Type 2 Diabetes." *Medical Decision Making (Suppl)*, 2002: S80-S91.

Eddy, David M, and L Schlessinger. "Archimedes: A trial-validated model of diabetes." *Diabetes Care*, 2003: 3093-3101.

Eddy, David M, and L Schlessinger. "Validation of the Archimedes Diabetes Model." *Diabetes Care*, 2003: 3102-3110.

Eliasson, B, J Cederholm, P Nilsson, and S Gudbjornsdottir. "The gap between guidelines and reality: Type 2 diabetes in a national diabetes register 1996-2003." *Diabetic Medicine*, 2005: 1420-1426.

Esposito, Katherine, et al. "Addition of Netural Protamine Lispro Insulin or Insulin Glargine to Oral Type 2 Diabetes Regimens for Patients with Suboptimal Glycemic Control." *Annals of Internal Medicine* 149 (2008): 531-539.

Evans, William E, and Julie A Johnson. "Pharmacogenomics: The Inherited Basis for Interindividual Differences in Drug Response." *Annu. Rev. Genomics Hum. Genet.* (Annu Rev Genomics Hum Genet) 2 (2001): 9-39.

Frick, M H, et al. "Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia. Safety of treatment, changes in risk factors, and incidence of coronary heart disease." *NEJM* 317, no. 20 (1987): 1237-1245.

Gaede, Peter, and Oluf Pedersen. "Intensive Integrated Therapy of Type 2 Dabetes: Implications for Long-Term Prognosis." *Diabetes*, 2004: S39-S47.

Gaede, Peter, Pernille Vedel, Hans-Henrik Parving, and Oluf Pedersen. "Intensified Multifactorial Intervention in Patients with Type 2 Diabetes Mellitus and Microalbuminuria: the Steno type 2 ranomised study." *The Lancet* 353 (1999): 617-22.

Garber, Alan J, Theodore G Duncan, Anita M Goodman, Donna J Mills, and Jane L Rohlf. "Efficacy of Metformin in Type 2 Diabetes: Results of a Double_Blind, Placebo-controlled, Dose-Response Trial." *The Am. Journ. Of Medicine* 102 (1997): 491-497.

Gigerenzer, Gerd, and Ulrich Hoffrage. "How to Improve Bayesian Reasoning Without Instruction: Frequency Formats." *Psychological Review*, 1995: 684-704.

Gilmer, T P, et al. "A Diabetes Simulated Physician Learning Program Improved Glucose Control and is Highly Cost-Effective." *Poster at the ADA's 70th Scientific Sessions.* Orlando, FL, 2010.

Group, ADVANCE Collaborative. "Effects of a fixed combination of perindopril and indapamide on macrovascular and microvascular outomces in patients with type 2

diabetes mellitus (the ADVANCE trial): a randomised controlled trial." *Lancet* 370 (2007): 829-40.

Group, Heart Protection Study Collaborative. "MRC/BHF Heart Protection Study of Choleterol-lowering with Simvastatin in 5963 People with Diabetes: A Randomised Placebo-controlled Trial." *The Lancet* 361 (2003): 2005-16.

Group, The ACCORD. "Effects of Intensive Glucose Lowering in Type 2 Diabetes." *The New England Journal of Medicine* 358 (2008): 2545-59.

Group, The ADVANCE Collaborative. "Intensive Blood Glucose Control and Vascular Outcomes in Patients with Type 2 Diabetes." *N Engl J Med* 358, no. 24 (2008): 2560-72.

Group, The CDC Cost Effectiveness. "Cost-effectiveness of Intensive Glycemic Control, Intesnisifed Hypertension Control, and Serum Cholesterol Level Reduction for Type 2 Diabetes." *JAMA*, 2542-2551: 2002.

Group, The Mount Hood 4 Modeling. "Computer Modeling of Diabetes and Its Complications." *Diabetes Care* 30, no. 6 (2007).

Guyatt, Gordon, David Sackett, Wayne Taylor, John Chong, Robin Roberts, and Stewart Pugsley. "Determining Optimal Therapy -- Randomized Trials in Individual Patients." *The New England Journal of Medicine* 314 (1986): 889-892.

Guyatt, Gordon, et al. "A Clinician's Guide for Conducting Randomized Trials in Individual Patients." *Canadian Medical Association Journal* 139 (1988): 497-503.

Hammond, K R. *Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice.* New York: Oxford University Press, Inc, 1996.

Hansson, L, et al. "Randomised trial of effects of calcium antagonists compared with diuretics and B-blockers on cardiovascular morbidity and mortality in hypertension: the Nordic Diltiazem (NORDIL) study." *Lancet*, 2000: 359-65.

Herman, W H, et al. "The Cost-Effenctivenss of Lifestyle Modification or Metformin in Preventing Type 2 Diabetes in Adults with Impaired Glucose Tolerance." *Ann Intern Med*, 2005: 323-332.

Hoerger, T J, R Harris, K A Hicks, K Donahue, S Sorensen, and M Engelgau. "Screening for Type 2 Diabetes Mellitus: A Cost-Effectiveness Analysis." *Ann Intern Med*, 2004: 689-699.

Holman, Rury R, Sanjoy K Paul, M Angelyn Bethel, David R Matthews, and H Andrew W Neil. "10-Year Follow-up of Intensive Glucose Control in Type 2 Diabetes." *New England Journal of Medicine* 359 (2008).

Holman, Rury R, Sanjoy K Paul, M Angelyn Bethel, H Andrew W Neil, and David R Matthews. "Long-Term Follow-up after Tight Control of Blood Pressure in Type 2 Diabetes." *New England Journal of Medicine* 359 (2008).

Hu, Frank B, Meir J Stampfer, Steven M Haffner, Caren G Solomon, Walter C Willett, and JoAnn E Manson. "Elevated Risk of Cardiovascular Disease Prior to Clinical Diagnosis of Type 2 Diabetes." *Diabetes Care*, 2002: 1129-1134.

Huttunen, J K, et al. "The Helsinki Heart Study: Central Findings and Clinical Implications." *Annals of Medicine* 23 (1991): 155-159.

Inzucchi, Silvio E. "Oral Antihyperglycemic Therapy for Type 2 Diabetes." *JAMA*, 2002: 360-372.

Issenberg, S Barry, et al. "Simulation Technology for Health Care Professional Skills Training and Assessment." *JAMA*, 1999: 861-866.

James, F. "Monte Carlo theory and practise." *Reports on Progress in Physics*, 1980: 1145-1190.

Jones, P, S Kafonek, I Laurora, and D Hunningshake. "Comparative Dose Efficacy Study of Atorvastatin Versus Simvastatin, Pravastatin, Lovastatin, and Fluvastatin in Ptients with Hpyercholesterolemia (The CURVES Study)." *Am J Cardiol* 81 (1998): 582-587.

Julius, S, et al. "Outcomes in hypertensive patients at high cardiovascular risk treated with regimens based on valsartan or amlodipine: a VALUE randomised trial." *The Lancet* 363, no. 9426 (2004): 2022-2031.

Katon, Wayne J, et al. "The Association of Comorbid Depression with Mortality in Patients with Type 2 Diabetes." *Diabetes Care*, 2005: 2668-2672.

Kornitzer, M, M Dramaix, M D Vandenbroek, L Everaert, and C Gerlinger. "Efficacy and tolerance of 200mg micronised fenofibrate administered over a 6-month period in hyperlipidaemic patients: an open Belgian multicenter study." *Atherosclerosis* 110 (suppl) (1994): S49-S54.

Lachin, John M, Saul Genuth, David M Nathan, Bernard Zinman, and Brandy N Rutledge. "Effect of Glycemic Exposure on the Risk of Microvascular Complications in the Diabetes Control and Complications Trial -- Revisited." *Diabetes*, 2008: 995-1001.

Laville, Frederic. "Foundations of Procedural Rationality: Cognitive Limits and Decision Processes." *Economica and Philosophy* 16 (2000): 117-138.

Lebovitz HE, Dole JF, Patwardhan R, Rappaport EB, Freed MI, The Rosigltazone Clinical Trials Study Group., et al. "Rosiglitazone Monotherapy Is Effective in Patients with Type 2 Diabetes." *J Clin Endocrinol Metab* 86 (2001): 280-288.

Levy, J, A Atkinson, P Bell, D McCance, and D Hadden. "Beta-cell Deterioration Determines the onset and rate of Progression of Secondary Dietary Failure in Type 2 Diabetes Mellitus: The 10-year follow-up of the Belfast Diet Study." *Diabetic Medicine* 15, no. 4 (2004): 290-296.

Mahon, Jeffrey, Andreas Laupacis, Allan Donner, and Thomas Wood. "Randomised Study of N of 1 Trials Versus Standard Practice." *British Medical Journal* 312 (1996): 1069-1074.

Mancia, G, G Parati, G Pomidossi, G Grassi, R Casadei, and A Zanchetti. "Alerting reaction and rise in blood pressure during measurement by physician and nurse." *Hypertension* 9 (1987): 209-215.

March, James G, Lee S. Sproull, and Michal Tamuz. "Learning from Samples of One or Fewer." *Organization Science* 2, no. 1 (1991): 1-13.

Marr, David. "Artificial Intelligence -- A Personal View." *Artificial Intelligence* 9 (1977): 37-48.

—. *Vision: A Computational Investigation into the Human Respresentation and Processing of Visual Information.* Edited by Judith Wilson. San Francisco: W.H. Freeman and Company, 1982.

Materson, B J, et al. "Single-Drug Therapy for Hypretension in Men -- A Comparison of Six Antihypertensive Agents with Placebo." *NE J of Medicine* 328 (1993): 914-921.

Mazze, R, E Strock, and G Simonson. "Staged Diabetes Management: Detection and Treatment." Minneapolis, MN, 2001.

McCabe, R M, et al. "Validation of the SimCare Model: A Computational Model of Individual Patients with Type 2 Diabetes." *Poster at the ADA's 70th Scientific Sessions.* Orlando, FL, 2010.

McCabe, Ryan M, et al. "Using Data Mining to Predict Errors in Chronic Disease Care." In *Advances in Patient Safety: New Directions and Alternative Approaches*, by Agency for Healthcare Research and Quality, 349-367. 2008.

Meyer, G, et al. "A Machine Learning Approach to Improving Process Control." *Proceedings of the 19th Workshop on Information Technologies and Systems.* Phoenix, AZ, 2009.

—. "Towards Lower Macrovascular Risk in Diabetes Patients: A Simulation-Based Evaluation of Prioritization Strategies." *ADA 70th Scientific Sessions.* Orlando, FL, 2010.

Mikulecky, Donald C. "Robert Rosen: The Well-Posed Question and its Answer -- Why are organisms different from machines?" *Systems Research and Behavioral Science* 7 (2000): 419-432.

Mitchell, Jason P. "Mentalizing and Marr: An information processing approach to the study of social cognition." *Brain Research* 1079 (2006): 66-75.

Mogensen, C E, et al. "Randomised controlled trial of dual blockade of renin-angiotensin system in patients with hpyertension, microalbuminuria, and non-insulin dependnet diabetes: the candesartan and lisinopril microalbuminuria (CALM) study." *British Medical Journal* 321 (2000): 1440-1444.

Muller, E, et al. "Development and validation of the Economic Assessment of Glycemic Control and long-term effects of diabetes (EAGLE) model." *Diabetes Technol Ther*, 2006: 219-236.

Muller, E, S Maxion-Bergemann, B Bolinder, R A Gerber, and R Bergemann. "EAGLE--Economic Assessment of Glycemic Control and Longterm Effects: a computer simulation model for diabetes mellitus type 1 and type 2 (Abstract)." *Diabetologia*, 2004: A355.

Muller, E, S Maxion-Bergemann, D Gultyaev, S Walzer, and R Bergemann. "EAGLE diabetes model: basic features and internal validation of simulating long-term diabetic outcomes and related costs (Abstract)." *Value Health*, 2004: 745.

Nathan, D. M., et al. "Medical Managment of Hyperglycaemia in Type 2 Diabetes Mellitus: A consensus algorithm for the initiation and adjustment of therapy." *Diabetologia* 52 (2009): 17-30.

O'Connor, P J, et al. "Simulated Physician Learning Intervention to Improve Safety and Quality of Diabetes Care: A Randomized Trial." *Diabetes Care*, 2009: 585-590.

O'Connor, P J, J M Sperl-Hillen, P Johnson, P J Rush, and A L Crain. "Customized Feedback to Patients and Providers Failed to Improve Safety or Quality of Diabetes Care: A Randomized Trial." *Diabetes Care*, 2009: 1158-1163.

Organization, Collaborating Healthcare.

Palmer, A J, et al. "The CORE Diabetes Model: Projecting long-term clinical outcomes, costs and cost-effectiveness of interventions in diabetes mellitus (types 1 and 2) to support clinical and reimbursement decision-making." *Curr Med Research And OPin(Suppl 1)*, 2004: S5-S26.

Palmer, A J, et al. "Validation of the CORE Diabetes Model against epidemiological and clinical studies." *Curr Med Research and Opin(Suppl 1)*, 2004: S27-S40.

Pew, Richard W, and Anne S Mavor. *Modeling Human and Organizational Behavior: Application to Military Simulations.* Washington, D.C.: National Academy Press, 1998.

Phillipov, G, and P J Phillips. "Components of Total Measurement Error for Hemoglobin A1c Determination Clinical Chemistry." *Clinical Chemistry* 47, no. 10 (2001): 1851-1853.

Rahman, Mahboob, et al. "Renal Outcomes in High-Risk Hypertensive Patients Treated With an Angiotensin-Converting Enzyme Inhibitor or a Calcium Channel Blocker vs a Diuretic." *Archives of Internal Medicine* 165 (2005): 936-946.

Ramsey, G W, P E Johnson, G Adomavicius, R M McCabe, M Elidrisi, and P J O'Connor. "Improving Chronic Disease Care Using Predictive Modeling and Data Mining." *3rd INFORMS Workshop on Data Mining and Health Informatics.* Washington, D.C., 2008.

Ramsey, G W, P E Johnson, P J O'Connor, J M Sperl-Hillen, and W A Rush. "Computational Models for Investigating Success and Failure in Treating Patients withType 2 Diabetes." *Proceedings of the 5th INFORMS Workshop on Data Mining and Health Informatics.* 2010.

—. "Using Functional Data Analysis to Identify Physician Strategies which Lead to Better Type 2 Dabetes Patient Outcomes." *Proceedings from the 1st ACM International Health Informatics Symposium.* Arlington, VA, 2010.

Ramsey, G W, P E Johnson, P J O'Connor, J M Sperl-Hillen, W A Rush, and G Biltz. "Identifying Physician Decision Strategies for Treating Patients with Type 2 Diabetes." *Poster at the ADA's 70th Scientific Sessions.* Orlando, FL, 2010.

Rickheim, P L, J L Flader, T W Weaver, and D M Kendall. "Assessment of Group Versus Individual Diabetes Education." *Diabetes Care* 25, no. 2 (2002): 269-274.

Roese, Neal J. "Counterfactual Thinking." *Psyholocial Bulletin* 121, no. 1 (1997): 133-148.

Rohlfing, C L, H M Wiedmeyer, R R Little, J D England, A Tennill, and D E Goldstein. "Defining the Relationship Between Plasma Glucose and HbA1c." *Diabetes Care* 25, no. 2 (2002): 275-278.

Rosenstock, J, S L Schwartz, C M Clark, G D Park, D W Donley, and M B Edwards. "Basal Insulin Therapy in Type 2 Diabetes:28-week comparison of insulin glargine(HOE 901) and NPH insulin." *Diabetes Care* 24 (2001): 631-636.

Rubin, Robert R. "Adherence to pharmacologic therapy in patients with type 2 diabetes mellitus." *The American Journal of Medicine*, 2005: 27-34.

Sackett, David L, William M C Rosenberg, J A Muir Gray, R Brian Haynes, and W Scott Richardson. "Evidence based medicine: what it is and what it isn't." *BMJ*, 1996.

Sacks, David B. "Translating Hemoglobin A1c into Average Blood Glucose: Implications for Clinical Chemistry." *Clinical Chemistry*, 2008: 1756-1758.

Sacks, Frank M, et al. "The Effect of Pravastatin on Coronary Events After Myocardial Infarction in Patients with Average Cholesterol Levels." *The New England Journal of Medicine* 335, no. 14 (1996): 1001-9.

Sargent, Robert G. "Validation and Verification of Simulation Models." Edited by R G Ingalls, M D Rossetti, J S Smith and B A Peters. *Porceedings of the 2004 Winter Simulation Conference.* 2004. 17-28.

Saudek, Christopher D, Rachel L Derr, and Rita R Kalyani. "Assessing Glycemia in Diabetes Using Self-monitoring Blood Glucose and Hemoglobin A1c." *JAMA*, 2006: 1688-1697.

Sperl-Hillen, J M, et al. "A New Approach to CME? A Simulated Physician Learning Program Improves Glucose Control in Adults with Diabetes." *Poster at the ADA's 70th Scientific Sessions.* Orlando, FL, 2010.

Sperl-Hillen, J M, et al. "A Simulated Physician Learning Program Improves Glucose Control in Adults with Diabetes." *Diabetes Care*, 2010: 1727-1733.

—. "Personalized Physician Learning Intervention Improved Glocose Control in Adults with Diabetes." *Diabetes.* 2009.

Stamler, J, O Vaccaro, J D Neaton, and D Wentworth. "Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial." *Diabetes Care*, 1993: 434-444.

Standl, E, B Balletshofer, B Dahl, B Weichenhain, H Stiegler, and A Hormann. "Predictors of 10-year macrovascular and overall mortality in patients with NIDDM: the Munich General Practitioner Project." *Diabetologia*, 1996: 1540-1545.

Stevens, R J, V Kothari, A I Adler, I M Stratton, and R R Holman. "The U KPDS risk engine: a model for the risk of coronary heart disease in type II diabetes (UKPDS 56)." *Clinical Science*, 2001: 671-679.

Tetlock, Philip E. *Expert Political Judgment: How good is it? How can we know?* Princeton: Princeton University Press, 2005.

Tran, David V, Tammy L Hofer, Terrence Lee, and George S Cembrowski. "Unique Approach to Derivation of Random Error in Laboratory Assays: Application to Glycohemoglobin Testing Demonstrates Poor Clinical Performance for Immunochemistry Assays." *Diabetes Technology & Therapeutics*, 2003: 975-978.

Tseng, C L, et al. "Seasonal Patterns in Monthly Hemoglobin A1c Values." *Am J Epidemiol* 161 (2005): 565-574.

Turner, R C, H Millns, H A Neil, I M Stratton, S E Manley, and D R Matthews. "Risk factors for coronary artery disease in non-insulin dependent diabetes mellitus: United Kingdom Prospective Diabetes Study (UKPDS: 23)." *BMJ*, 1998: 823-838.

Tversky, A, and D Kahneman. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185 (1974): 1124-1131.

Wannamethee, S G, and A G Shaper. "Weight Change and Duration of Overweight and Obesity in the Incidence of Type 2 Diabetes." *Diabetes Care* 22 (1999): 1266-1272.

Wen, S W, M S Kramer, J Hoey, J A Hanley, and R H Usher. "Terminal digit preference, random error, and bias in routine clinical measurement of blood pressure." *J Clin Epidemiol* 46, no. 10 (1993): 1187-93.

Whooley, Mary A, and Gregory E Simon. "Managing Depression in Medical Outpatients." *New England Journal of Medicine* 343 (2000): 1942-1950.

Wilkonson, G R. "Drug Metabolism and Variability among Patients in Drug Response." *N Engl J Med* 352 (2005): 2211-2221.

Appendix Overview

This collection of appendices exists to give the reader of this thesis access to important detail that supports the work of this thesis but would not appropriately fit into the body of writing.

Appendix A: Dose Response Curves

The tables of this appendix detail the full, dose-by-dose definitions of every dose response curve used in the SimCare model.

Appendix B: University of Minnesota SimCare Version

After the initial implementation of the SimCare model in previous work, two versions of the software branched to accommodate overlapping research teams and priorities. One version went to a collaborating healthcare organization and one version remained with the University of Minnesota. This appendix details any aspects of the University of Minnesota version that have been added for research purposes that may not have been added to the other version.

Appendix C: Inactive Factors

This appendix details a number of factors that were initially suggested for inclusion but were never fully defined or implemented.

Appendix D: History of Insulin Effects Changes

The way insulin works in the SimCare model is a technical topic, with great detail that would be of interest to certain audiences. Over time, the definition of this process has been refined. The history of these changes has been maintained in this appendix.

Appendix E: N=1 Simulation Error Data

Due to the nature of N=1 research, valuable information is contained in individual experiments, especially cases of error. The full set of data for the error cases are listed in this appendix.

Appendix F: Fourth Mt. Hood Challenge Model Review

The Mt. Hood Challenge was a series of conferences held for computational modelers of type 1 and type 2 diabetes. The fourth challenge in particular held a comparison competition for a variety of available models at the time. This appendix is a condensed review of the models that participated in that challenge and is based on the available literature at the time following the fourth challenge.

Appendix G: A Software Implementation of the SimCare Patient Model

An overview of the related software systems and Physician Process Models is given to create a working picture of the ecology in which the SimCare model exists.

Appendix A: Dose Response Curves

Table 1A. Dose response curves. This table details the effect per unit (mg except for insulin) of each drug currently modeled in SimCare. For example, the effect per mg of Metformin between 0 and 500mg is a .0006 drop of A1c; so the total effect for titrating from 0 to 500 mg is (500 x .0006) = .3; a .3% drop in A1c. Titrating from 500mg to 1,000 mg will result in an additional drop of ((1000 – 500) x .001) = .5; a ½ percent drop in A1c.

Medication	Affected variable	**In bold: Dose.** Below dose, marginal effect per mg (or unit for insulin) on affected variable						Max Effect
Blood Glucose Medications								
Metformin	**Dose mg**	**500**	**1000**	**1500**	**2000**			
	SMBG (mg/dL)	-40.0	-13.3	-22.2	-13.3			-88.8
	A1c (%)	-0.9	-0.3	-0.5	-0.3			-2.0
Glipizide	**Dose mg**	**5**	**10**	**20**				
	SMBG (mg/dL)	-35.52	-17.76	-13.32				-66.6
	A1c (%)	-0.8	-0.4	-0.3				-1.5
Glyburide	**Dose mg**	**2.5**	**5**	**10**				
	SMBG (mg/dL)	-44.4	-8.88	-13.32				-66.6
	A1c (%)	-1	-0.2	-0.3				-1.5
Glimeperide	**Dose mg**	**2**	**4**	**8**				
	SMBG (mg/dL)	-44.4	-8.88	-13.32				-66.6
	A1c (%)	-1	-0.2	-0.3				-1.5
Pioglitazone	**Dose mg**	**15**	**30**	**45**				
	SMBG (mg/dL)	-44.4	-4.4	-22.2				-71.0
	A1c (%)	-1.0	-0.1	-0.5				-1.6
Rosiglitazone	**Dose mg**	**2**	**4**	**8**				
	SMBG (mg/dL)	-53.3	-8.9	-4.4				-66.6
	A1c (%)	-1.2	-0.2	-0.1				-1.5
Lispro BMI 25	**Dose u/kg/d**	**0.5**	**1**	**2**	**3**			
	SMBG (mg/dL)	-73.396	-36.698	-24.465	-18.349			-97.861
	A1c (%)	-0.826	-0.41327	-0.55102	-0.4132			-2.2
Glargine BMI 25	**Dose u/kg/d**	**0.5**	**1**	**2**	**3**			
	SMBG	-186.6	-93.3	-62.2	-46.65			-248.8
	A1c (%)	-2.1013	-1.05068	-1.4009	-1.0506			-5.6
Blood Pressure Medications								
Amlodipine	**Dose mg**	**2.5**	**5**	**10**				
	SBP (mmHg)	-5	-5	-5				-15.0
Atenolol	**Dose mg**	**25**	**50**	**100**				
	SBP (mmHg)	-5	-5	-5				-15.0
Lisinopril	**Dose mg**	**5**	**10**	**20**	**40**	**60**	**80**	
	SBP (mmHg)	-5	-5	-5	-5	-5	-5	-30.0
	Creatinine	0.3	0.03	0.06	0.11			0.5
Valsartan	**Dose mg**	**20**	**40**	**80**	**160**	**240**	**320**	
	SBP (mmHg)	-5	-5	-5	-5	-5	-5	-30.0
	Creatinine	0.3	0.03	0.06	0.11			0.5
Hydrochlorothiazide	**Dose mg**	**12.5**	**25**					
	SBP (mmHg)	-7.5	-6.5					-14.0
Cardizem	**Dose mg**	**60**	**120**	**180**	**240**			

	SBP (mmHg)	-5	-5	-5	-5			-20.0
Lipid Medications								
Simvastatin	**Dose mg**	**5**	**10**	**20**	**40**	**80**		
	LDL (% mg/dL)	-20	-5	-10	-10	-10		-55.0
	HDL (mg/dL)	2	2	2	2			+8.0
	Triglycerides (% mg/dL)	-3	-3	-6	-13			-25.0
Atorvastatin	**Dose mg**	**5**	**10**	**20**	**40**	**80**		
	LDL (% mg/dL)	-25	-5	-10	-10	-10		-60.0
	HDL (mg/dL)	3	2	2	2			+9.0
	Triglycerides (% mg/dL)	-4	-4	-7	-16			-31.0
Fenofibrate	**Dose mg**	**67**	**134**	**200**				
	LDL (% mg/dL)	-5						-5.0
	HDL (mg/dL)	5	5	5				+15.0
	Triglycerides (mg/dL)	-10	-10	-20				-40.0
Gemfibrazol	**Dose mg**	**1200**						
	HDL (mg/dL)	15						+15.0
	Triglycerides (% mg/dL)	-35						-35.0
Depression Medication								
Zoloft	**Dose mg**	**50**	**100**	**150**				
	Depression (%)	-30	-30	-30				-90

Appendix B: University of Minnesota SimCare Version

The SimCare Patient Model is constantly under analysis and development as additional information emerges about existing constraints in the model, or as new constraints are considered for inclusion. Initially, both implementations of the SimCare model were identical; one was used by the University of Minnesota and the other was used by the collaborating health organization. Over time, additional pathways were included for research purposes in the University of Minnesota implementation.

Pathways

- Seasonal variation in adherence: patient adherence can vary between -6.5% to +6.5% for a given patient depending on the season of the year (Tseng, et al. 2005).

- Inter-patient variation in SMBG: Two patients with the same A1c and ambient (i.e., average) SMBG reading may have vastly different SMBG sample variation throughout the day (Derr, et al. 2003). Specific SMBG values are generated by the model from the ambient SMBG for 5 different readings per day (breakfast, lunch, dinner, bedtime, 3am). The variation in these readings is modeled with a measurement fluctuation of -20.5% to +20.5%.

Appendix C: Inactive Factors

There exists code for the following variables and pathways but they are not active in the current version of the SimCare model.

Patient Health Characteristics

- Cardiovascular and microvascular risk (these is instead computed using the UKPDS Risk and Outcome models)
- Age (this is tracked for risk purposes when the patient population is created, but does not serve any purpose within the SimCare model)
- Smoking (this is tracked for risk purposes when the patient population is created, but does not serve any purpose within the SimCare model)
- Diastolic blood pressure (fixed as systolic divided by 1.7)
- Alcohol
- Stress

Pathways

- Aspirin (variables have been created but no dose response curve is set up)

Referrals

- Cardiologist
- Endocrinologist

Appendix D: History of Insulin Effects Changes

Before change:

$$DoseResponse_{Insulin,A1c}(x) = \begin{cases} 5(x_t - x_{(t-1)});\ 0 < x \leq .2 \\ 3.33(x_t - x_{(t-1)}) + 1;\ .2 < x \leq .5 \\ 3.6(x_t - x_{(t-1)}) + 2;.5 < x \leq 1 \\ 2.4(x_t - x_{(t-1)}) + 3.8; 1 < x \leq 2 \\ 6.2; 2 < x \end{cases}$$

After change:

$$DoseResponse_{Insulin,A1c}(x) = \begin{cases} 5(x_t - x_{(t-1)});\ 0 < x \leq .2 \\ 3.33(x_t - x_{(t-1)}) + 1;\ .2 < x \leq .5 \\ 3.6(x_t - x_{(t-1)}) + 2;.5 < x \leq 1 \\ 2.4(x_t - x_{(t-1)}) + 3.8; 1 < x \leq 3 \\ 8.6; 3 < x \end{cases}$$

Change of Insulin Effect on Blood Glucose

Insulin L

$$Insulin_L Effect_{BG} = -1244FD$$

$$F = \begin{pmatrix} 0 & 0 & .045 & .25 \\ 0 & 0 & 0 & .045 \\ .25 & 0 & 0 & 0 \\ .045 & .25 & 0 & 0 \\ 0 & .045 & .25 & 0 \end{pmatrix}, D = \begin{pmatrix} Dose_{Breakfast} \\ Dose_{Lunch} \\ Dose_{Supper} \\ Dose_{Bedtime} \end{pmatrix}$$

$Dose_{Breakfast} + Dose_{Lunch} + Dose_{Supper} + Dose_{Bedtime}$ = Daily insulin dose d

Insulin L effect on daily average blood glucose:

$-1244(Dose_{Bedtime}*.25 + .045*Dose_{Supper} + Dose_{Bedtime}*.045 + Dose_{Breakfast}*.25 + Dose_{Lunch}*.25 + Dose_{Breakfast}*.045 + Dose_{Supper}*.25 + Dose_{Lunch}*.045)/5$
$= -1244(.295(Dose_{Breakfast} + Dose_{Lunch} + Dose_{Supper} + Dose_{Bedtime}))$
$= -1244(.295d)/5 = -73.396d$

Insulin G

Insulin G effect on Blood Glucose before change:

$$Insulin_L Effect_{BG} = -1244FD$$

$$F = \begin{pmatrix} .15 & .15 & .175 & .15 \\ .125 & .15 & .15 & .175 \\ .15 & .125 & .15 & .15 \\ .175 & .15 & .125 & .15 \\ .15 & .175 & .125 & .15 \end{pmatrix}, D = \begin{pmatrix} Dose_{Breakfast} \\ Dose_{Lunch} \\ Dose_{Supper} \\ Dose_{Bedtime} \end{pmatrix}$$

$Dose_{Breakfast} + Dose_{Lunch} + Dose_{Supper} + Dose_{Bedtime}$ = Daily insulin dose d

Insulin G effect on daily average blood glucose before patch:

$-1244(.75(Dose_{Breakfast} + Dose_{Lunch} + Dose_{Bedtime} + Dose_{Supper}))/5 = -1244(.75d)/5 = -186.6d$

Insulin G effect on Blood Glucose after change

$$F = \begin{pmatrix} .15 & .15 & .15 & .15 \\ .175 & .175 & .175 & .175 \\ .15 & .15 & .15 & .15 \\ .15 & .15 & .15 & .15 \\ .125 & .125 & .125 & .125 \end{pmatrix}, D = \begin{pmatrix} Dose_{Breakfast} \\ Dose_{Lunch} \\ Dose_{Supper} \\ Dose_{Bedtime} \end{pmatrix}$$

Insulin G effect on daily average blood glucose:

$-1244((.15+.175+.15+.15+.125)(Dose_{Breakfast} + Dose_{Lunch} + Dose_{Bedtime} + Dose_{Supper}))/5 = -1244(.75d)/5 = -186.6d$

The blood glucose insulin G change only changed the within-day blood glucose values that would be reported through SMBGs, but did not have an effect on the daily average.

Appendix E: N=1 Simulation Error Data

Appendix A contains the records – encounter by encounter – of each simulated patient in the N=1 set of experiments who ended above or below the experimental range.

Tables 1E-5E: "Missed High." The following tables show encounter by encounter detail for the simulation experiments that ended in the error of the real patient's observed A1c level being higher than the SimCare upper bound. **Tables 6E-13E:** "Missed Low." These tables show the same detail for simulation experiments that ended in the error of the real patient's observed A1c level being lower than the SimCare lower bound. Bold values for the "Clinical A1c" reflect an updated A1c value in the patient record, otherwise, values were carried forward. SimCare A1c values always reflect the current state of the simulated patient.

"Missed High"

1E

Encounter Num	1	2	3	4
Encounter Date	0	99	107	163
Metformin	500	500	500	500
Glipizide	10	10	10	10
Pioglitazone	0	0	0	0
Insulin	0	0	0	0
Sim A1c Upper	8.0	8.4	8.5	8.7
Clinical A1c	**8.0**	**8.9**	8.9	8.9
Sim A1c Lower	8.0	8.0	8.0	7.8

2E

Encounter Num	1	2	3	4	5	6
Encounter Date	0	9	28	35	119	133
Metformin	0	1000	1000	1000	1000	1000
Glipizide	0	10	10	10	10	10
Pioglitazone	0	15	15	15	15	15
Insulin	0	0	0	0	0	0
Sim A1c Upper	6.5	6.5	6.4	6.2	6.3	6.4
Clinical A1c	**6.5**	6.5	**14.6**	14.6	14.6	**6.5**
Sim A1c Lower	6.5	6.4	3.8	3.2	3.2	3.2

3E

Encounter Num	1	2	3	4	5	6	7	8
Encounter Date	0	48	59	78	92	97	99	185
Metformin	500	500	500	500	500	500	500	500
Glipizide	5	5	5	5	5	5	5	5
Pioglitazone	0	0	0	0	0	0	0	0
Insulin	0	0	0	0	0	0	0	0
Sim A1c Upper	8.2	8.4	8.5	8.5	8.6	8.6	8.6	9.0
Clinical A1c	**8.2**	**9.1**	9.1	9.1	9.1	9.1	9.1	9.1
Sim A1c Lower	8.2	8.2	8.2	8.1	8.0	8.0	8.0	7.7

4E

Encounter Num	1	2	3	4
Encounter Date	0	3	76	81
Metformin	500	500	500	500
Glipizide	0	0	0	0
Pioglitazone	0	0	0	0
Insulin	0	0	0	0
Sim A1c Upper	6.7	6.7	7.0	7.1
Clinical A1c	**6.7**	6.7	**7.3**	7.3
Sim A1c Lower	6.7	6.7	6.4	6.4

5E

Encounter Num	1	2	3
Encounter Date	0	10	108
Metformin	1000	1000	1000
Glipizide	0	0	10
Pioglitazone	0	0	0
Insulin	0	0	0
Sim A1c Upper	6.5	6.5	7.0
Clinical A1c	**6.5**	**6.8**	**7.5**
Sim A1c Lower	6.5	6.5	6.3

"Missed Low"

6E

Encounter Num	1	2	3	4
Encounter Date	0	106	155	291
Metformin	0	500	500	500
Glipizide	0	10	10	10
Pioglitazone	0	0	0	0
Insulin	0	0	0	0
Sim A1c Upper	9.1	10.0	9.7	10.2
Clinical A1c	**9.1**	9.1	**5.4**	5.4
Sim A1c Lower	9.1	8.8	6.3	6.1

7E

Encounter Num	1	2	3
Encounter Date	0	56	149
Metformin	0	500	500
Glipizide	0	0	0
Pioglitazone	0	0	0
Insulin	0	0	0
Sim A1c Upper	10.1	10.8	10.9
Clinical A1c	**10.1**	**8.4**	8.4
Sim A1c Lower	10.1	9.8	8.6

8E

Encounter Num	1	2	3	4
Encounter Date	0	103	222	315
Metformin	500	500	500	500
Glipizide	0	0	5	5
Pioglitazone	0	0	0	0
Insulin	0	0	0	0
Sim A1c Upper	8.0	8.5	9.0	9.2
Clinical A1c	**8.0**	8.0	**7.5**	**6.4**
Sim A1c Lower	8.0	8.0	8.0	7.0

9E

Encounter Num	1	2	3	4	5	6
Encounter Date	0	22	63	126	162	194
Metformin	0	0	500	500	500	500
Glipizide	0	0	0	0	0	0
Pioglitazone	0	0	0	0	0	0
Insulin	0	0	0	0	0	0
Sim A1c Upper	8.5	9.0	9.2	9.2	9.4	9.5
Clinical A1c	**8.5**	**7.1**	7.1	7.1	**6.7**	6.7
Sim A1c Lower	8.5	8.3	8.2	7.0	7.0	7.0

10E

Encounter Num	1	2	3	4	5	6	7	8	9
Encounter Date	0	42	62	73	154	172	234	291	299
Metformin	0	0	0	0	0	0	1000	1000	1000
Glipizide	0	10	10	10	10	10	10	10	10
Pioglitazone	0	0	0	0	0	15	15	15	15
Insulin	0	0	0	0	0	0	0	0	0
Sim A1c Upper	11.2	11.8	11.7	11.7	11.9	12.0	12.0	11.8	11.8
Clinical A1c	**11.2**	11.2	11.2	11.2	**7.6**	7.6	**5.7**	**6.0**	6.0
Sim A1c Lower	11.2	10.9	9.9	9.5	9.0	9.0	7.9	6.6	6.6

11E

Encounter Num	1	2	3	4	5	6	7
Encounter Date	0	23	63	111	139	160	258
Metformin	0	0	0	0	500	500	500
Glipizide	0	0	0	0	0	5	5
Pioglitazone	0	0	0	0	0	0	0
Insulin	0	0	0	0	0	5	7
Sim A1c Upper	15.1	15.6	15.8	16.0	16.1	16.1	16.2
Clinical A1c	**15.1**	15.1	**14.7**	14.7	14.7	**9.7**	9.7
Sim A1c Lower	15.1	14.9	14.8	14.8	14.8	14.0	11.8

12E

Encounter Num	1	2	3	4	5	6
Encounter Date	0	3	44	71	150	203
Metformin	0	500	500	500	500	500
Glipizide	0	0	0	0	0	0
Pioglitazone	0	0	0	0	0	0
Insulin	0	0	0	0	0	0
Sim A1c Upper	8.4	8.4	8.8	8.9	9.2	9.5
Clinical A1c	**8.4**	8.4	8.4	8.4	8.4	**6.7**
Sim A1c Lower	8.4	8.4	7.0	6.9	6.9	6.9

13E

Encounter Num	1	2	3	4	5	6	7	8	9
Encounter Date	0	1	2	3	13	34	55	66	73
Metformin	0	0	0	500	500	500	500	500	500
Glipizide	0	10	10	10	0	0	10	10	10
Pioglitazone	0	0	0	0	0	0	0	0	0
Insulin	0	0	0	0	0	0	0	0	0
Sim A1c Upper	9.5	9.5	9.5	9.5	9.6	9.7	9.9	9.8	9.8
Clinical A1c	**9.5**	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5
Sim A1c Lower	9.5	9.5	9.4	9.3	8.3	7.9	7.9	7.4	7.0

Encounter Num	10	11	12	13	14	15
Encounter Date	83	150	167	244	258	353
Metformin	500	500	500	500	500	500
Glipizide	10	10	10	10	10	10
Pioglitazone	0	0	0	0	0	0
Insulin	0	0	0	0	0	0
Sim A1c Upper	9.8	10.0	10.1	10.4	10.5	10.9
Clinical A1c	9.5	9.5	**6.2**	6.2	**5.7**	5.7
Sim A1c Lower	6.6	6.0	6.0	5.9	5.9	5.8

Appendix F: Fourth Mount Hood Challenge Model Review

The following diabetes models review is based in part on the most recent Mt. Hood Challenge in 2004 (T. M. Group 2007). The fourth Mt. Hood Challenge was a conference concerning diabetes models and validating those models for both Type 1 and 2 diabetes. This work only focuses on the models and validation exercises for type 2 diabetes models. The exercise for these models was to simulate a randomized clinical trial investigating the effects of 10mg Atorvastatin administered for approximately 4 years (CARDS) (Colhoun, et al. 2004).

The **CDC/RTI Type 2 Diabetes Progression Model** (Earnshaw, et al. 2002) (Hoerger, et al. 2004) (Herman, et al. 2005) (T. C. Group 2542-2551) uses a Markov model to simulate the progression of newly diagnosed patients simultaneously through five disease pathways: nephropathy, neuropathy, retinopathy, CHD, and stroke. This model places a greater emphasis on macrovascular events and incorporates the UKPDS Risk Engine to predict CVD events. However, the primary output of this model is QUALYs and the treatment costs it took to generate these QUALYs. Simulated cohorts are newly diagnosed patients in 10 year bands ranging from 25 to 94 years old and are further defined by sex, ethnicity, hypertension status, hypercholesterolemia status, and current smoking status. Cohorts are simulated through death or age 95. The transition probabilities (defined for each of the five pathways) depend on time since diagnosis, time between onset and diagnosis, age, sex, race, glycemic level, smoking, serum cholesterol level, and hypertension. Transition probabilities were defined based on UKPDS data,

Eastman model data and three other papers. This model has published several results but has not published a strictly validation piece of work.

Archimedes (Eddy and Schlessinger, Archimedes: A trial-validated model of diabetes 2003) (Eddy and Schlessinger, Validation of the Archimedes Diabetes Model 2003) is a model of the pathophysiology of several diseases at the level of the individual patient, including diabetes and its complications, coronary artery disease, congestive heart failure, and asthma. It uses a set of differential equations to simulate continuous time within a simulated patient (i.e. this model is "stateless"). The model was constructed in five main steps. The first step was to develop the conceptual description using "features" and dynamics between features. This was a qualitative step using expert advice and medical textbooks. The second step was to identify studies that pertain to the features chosen for the model. These studies were the epidemiological and clinical studies identified by experts as the foundations of their own understandings of the disease. The third step was to learn the parameters for each differential equation from the studies identified in the second step. The fourth step was to program these equations into Smalltalk, where individual and combination results were then validated against the studies to ensure proper implementation. The fifth step was to then simulate a clinical trial. A typical clinical trial takes 10 minutes to simulate using 250 PCs. The computational complexity of this model is a non-trivial feature of it.

Diabetes is simulated within the model at a very detailed representation. Each individual simulated patient possesses its own liver, pancreas, muscles, etc. Personal demographics

are also specified for each patient, e.g. sex, race, age, BMI, etc. Family history is included, and individual responsiveness to each treatment is sampled from a distribution and used as a multiplier (the same way it is in SimCare). Treatments for diabetes are then administered with the specific effects acting on each organ. For example, glyburide and Metformin both ultimately affect Fasting Plasma Glucose, but glyburide uses the pancreas pathway and metformin uses the liver pathway. In this way, the singular or combined forms of oral medications, insulins and diet and exercise are simulated. These treatments are administered by any one of more than a hundred highly specified treatment algorithms. An example of such a decision tree is "if the patient's LDL is > 180 and creatinine is < 2, then give lovastatin 80 mg. At 2 months, have the patient get a lipid panel and creatinine test. At that time if the LDL is not < 130 and the creatinine is still < 2, then switch to Simvastatin 80 mg…," etc. System resources, presumably meant to simulate the physical clinical environment are also modeled. Differences in personnel, facilities, equipment, and supplies combine to represent 37 different types of office visits. In this way, Archimedes does simulate clinical encounters.

An extensive amount of validation has been performed on Archimedes, conducting 74 exercises extracted from 18 separate trials. Ten of these trials were in no way used to build the model, and therefore, provided external validation. The other 8 were, in a piecewise way, used to build the model and provided an internal validation of it. An example of the validation (and building) method was that a relationship was learned from the 4-S study* between Simvastatin and rates of coronary artery occlusion. Once this relationship was modeled by equation, the equation was used in the composite model for

all subsequent trials and was not changed to fit any specific trial. To simulate a trial, large populations of patients are birthed by the model and allowed to grow up. From these huge populations (e.g. a thousand individuals may need to be simulated to find two who qualify for a trial), inclusion and exclusion criteria from the specific trial are used to select patients from the population. These groups are then verified to be within the distribution parameters specified by the trial. Over and under sampling is conducted until a fit is made (as in a real trial), and then randomization occurs.

For the 74 validation exercises, 71 were statistically identical to the real trial results. The correlation between model and trial for the trials not used to build the model was $r = 0.99$. The correlation between model and all trials was $r = 0.99$. Comparing the absolute differences in outcomes between the control and treatment groups, the correlation coefficient was $r = 0.97$. Archimedes is the most difficult to understand (least transparent) of all the models. It is the most vigorously validated relative to the other models, however it is not clear how vigorous that validation effort is relative to the massive increase in complexity of the model. No effort to validate individual patient treatment/disease paths or variation was noted. It is also not clear how much predicting event rates at the epidemiological level buys the validation process. Certainly, the repetition of the exercise gains it a credibility regarding its consistency (as well as the consistency of randomized clinical trials). It is however unclear how well this model would predict real patients in real clinical contexts.

The **Cardiff Diabetes Model** (Bagust, et al. 2001) is primarily a long-term economic model of the effects of treatments on cohorts of patients with Type 2 diabetes. This economic model (primarily based on the former Eastman model) was designed with three main objectives: 1) to characterize the lifetime experience of diabetes-related morbidities and use of health care resources among a typical population of patients diagnosed with Type 2 diabetes; 2) to estimate the difference in expected health outcomes and health costs attributable to Type 2 diabetes over a lifetime; 3) and to provide a basis for evaluation of future polices or interventions in the treatment of Type 2 diabetes from the perspective of a funder of health services. Cardiff is a deterministic Markov model, with separate transition probabilities specified for each of 15 age groups (banded by 5 years, after age 20) for each sex. These 30 models also run in parallel on patients without type 2 diabetes in order to describe marginal risks and costs attributable to diabetes. This provides estimates of the prevalence of all morbid states for any mixed age and sex cohort of newly diagnosed patients. This model does not contain "dose response curves" but models HbA1c as an estimate based on the effects of states of four therapies (diet and exercise, first and second-line oral hypoglycemic agents and insulin-based therapy). Each 30 instances of the model contains several linked modules, including: Complications (CHD and stroke, from South Glamorgan database), CVD Mortality (from Framingham and OPCS of UK), CVD (from Framingham), Retinopathy and macular edema (from WESDR or Eastman), Nephropathy (from WESDR or Eastman), Neuropathy, PVD and lower extremity amputation (WESDR or Eastman), HbA1c (from Hayward), and Costs (based on age-sex profiles from South Glamorgan database and GPMDP). This model represents individuals as seen from five year bands over

individual years of treatment. Specific treatments are not modeled by drug effects in general, but the effects (transition probabilities) are learned from databases. This model was originally designed to capture UK populations of patients.

The **Sheffield Diabetes Model** (T. M. Group 2007) has not been published other than its mention in the fourth Mt. Hood Challenge meeting. This model is apparently still under development but is designed to model the progression of type 2 diabetes. It uses five sub-models to represent comorbidity classifications of CHD, stroke, nephropathy, retinopathy, and neuropathy. The effects of therapy are simulated on the progression of the three main risk factors: A1c, lipids and blood pressure. This model clearly leans on UKPDS Risk Engine. The CHD and Stroke sub-models were derived from source data taken from UKPDS, British Heart Foundation statistics, the PRAIS-UK study, Minnesota population and NICE guidelines, the PROGRESS trial and the Heart Protection Study. The nephropathy, retinopathy and neuropathy sub-models are based primarily on the Eastman Model (plus data from Rochester Epidemiology Project, DCCT, and UKPDS 50).

The **UKPDS Outcomes Model** (Clarke, et al. 2004) was created using data from 3642 patients with type 2 diabetes over 14 years (1977 - 1991) of intensive or conventional control of blood glucose and BP. It predicts the occurrences of seven major complications related to diabetes: Ischemic heart disease (IHD), Myocardial infarction (MI), Congestive heart failure (CHF), Stroke, Amputation, Blindness, and Renal failure. Diabetes-related and other mortality are also simulated. This structure is mainly designed

to predict quality-adjusted life-years (QUALYs) for individual patients. By establishing first a baseline of risk factors (A1c, BP, smoking, etc), a probabilistic discrete-time illness-death model with annual cycles is used to simulate the occurrence of events. When a simulated patient dies, the QUALYs are tabulated for that iteration of the simulation, and the next iteration on the same patient is run. In this way, an average or expected QUALY is computed for each patient. Underlying changes in risk factors may be supplied by the user (based on other data) or are supplied by the model based on UKPDS average observations. (n.b., this is why A1c generally went up over time when we allowed the Outcomes Model to supply values -- A1c generally rose in the UKPDS population). This model consists of a family of independent Weibull hazard functions used to connect one risk parameter to any of the seven outcomes predicted. An equation was fitted and included in this family if it was found to be statistically significant in its relationship to the outcome. An example of this is that systolic blood pressure was found to be correlated to future occurrences of six of the seven outcomes but not blindness. Therefore, six different equations were fitted relating SBP to these outcomes and included in the model. Blindness as a past event is predictive of future amputation and renal failure. In this way, as events are simulated, QUALYs are estimated in a holistic way from disparate correlations. This overcomes the difficulty of learning a full joint distribution of risk factor/event outcome relationships. The UKPDS Outcomes model was validated against the UKPDS clinical trial data and predicted event rates in all cases within a 95% confidence interval.

The **UKPDS Risk Engine** (Stevens, et al. 2001) was designed to predict fatal and non-fatal CHD and fatal and non-fatal stroke in patients with type 2 diabetes. Built from UKPDS data, the Risk Engine specifically predicts first events and assumes no prior events. Data from 5102 patients followed for a median of 10.7 years were used to form the predictive models (excluding patients with events within the first 4 years of the study to remove selection bias). The model uses four parametric exponential equations to predict each of the four outcomes. Each equation uses all of the following individual patient inputs to create the estimation: age, sex, ethnicity, smoking status, A1c, SBP, HDL and total cholesterol. Each predictive model was fitted using a maximum likelihood estimation implemented by Newton-Raphson methods (a numerical method for empirically estimating functions by finding their roots). N.B. All of the Models mentioned in the Fourth Mt. Hood challenge used the UKPDS Risk engine to estimate events except UKPDS Outcomes Model, Archimedes and EAGLE. This is offered as a possible explanation for the overestimation of events trend among the models (UKPDS A1c trended upward).

The **Economic Assessment of Glycemic control and Long-term Effects (EAGLE)** (Muller, Maxion-Bergemann, et al., EAGLE diabetes model: basic features and internal validation of simulating long-term diabetic outcomes and related costs (Abstract) 2004) (Muller, Maxion-Bergemann and Bolinder, et al. 2004) (Muller, Maxion-Bergemann and Gultyaev, et al., Development and validation of the Economic Assessment of Glycemic Control and long-term effects of diabetes (EAGLE) model 2006) model is sparsely documented. The references sited in the Mt. Hood Challenge are only abstracts and

appear to not ever have been fully published. Another source from 2001 is contained within Mary Ann Liebert, Inc publishers and is not available without subscription. A company called Analytica may also be involved with this model and has been contacted. From what information on the model exists, it is a Markov model representing patient states as annual state transformations. Underlying A1c is modeled as a decrement from annual treatment regimens (state-based). States change annually based on the transition probabilities which are determined by the current state and the underlying A1c. Twenty outcomes (e.g. hypoglycemia, retinopathy, macular edema, end-stage renal disease, neuropathy, diabetic foot syndrome, MI and stroke) are projected for cohorts based on DCCT, UKPDS and WESDR (Wisconsin Epidemiological Study of Diabetic Retinopathy). The risks of these events occurring are defined by risk functions learned by regression analyses (using linear, exponential and quadratic) over these data. Events are assigned to individual simulated patients as they occur (using Monte Carlo sampling) and average event rates are then computed over the pre-specified iterations (e.g. 100) per cohort. Average event rate per defined cohort is the output of this model and is then used to fuel the economic arm of the model, which projects output data on costs, cost-consequence, quality of life, and cost-effectiveness of interventions. Definitions of these economic computations were not given.

The **CORE Diabetes Model** (Brandle and Herman 2004) (Palmer, Roze, et al., The CORE Diabetes Model: Projecting long-term clinical outcomes, costs and cost-effectiveness of interventions in diabetes mellitus (types 1 and 2) to support clinical and reimbursement decision-making 2004) (Palmer, Roze, et al., Validation of the CORE

Diabetes Model against epidemiological and clinical studies 2004) was developed by the Center for Outcomes research in Switzerland. This group has apparently either disbanded or withered since its two main researchers left (Palmer and Roze). Todd Gilmer emailed me two papers concerning this model, though it is unclear if he worked with this group directly or not. This model is another Markov, state-based epidemiological/policy assessment model designed to project the long-term clinical and health economic outcomes appropriate for patients with type 1 or 2 diabetes. The model is actually a composition of 14 independent sub-models (angina, MI, congestive heart failure, stroke, peripheral vascular disease, diabetic retinopathy, macula edema, cataract, hypoglycemia, ketoacidosis, lactic acidosis, nephropathy and end-stage renal disease, neuropathy, foot ulcer, and amputation) along with nonspecific mortality to simulate cohort risk factors and event complications. Treatment trees are described generally and based on probabilities as applied to individual patients, rather than as clinical treatment rules (e.g. if a patient is using inhaled insulin, then the CORE user specifies the probability that a treatment failure or side effect is experienced and that injected insulin is to be used instead). However, specific treatments are not simulated at the clinical encounter level. General, long-term trends in treatments are modeled as reactions to long-term trends in disease progression among randomized clinical trials.

Appendix G: A Software Implementation of the SimCare Patient Model

The technologies employed for the modeling efforts of this work include 1) a clinical patient model (SimCare), 2) a physician interface GUI, 3) a synthetic patient population generator 4) a clinical physician decision making model, and 5) a clinical physician/patient encounter simulator (Figure 1G). Each part of the modeling technology will be described in detail below. This work contributed significantly to the development of the clinical physician decision-making models and the clinical encounter simulator, however for the purposes of composing and organizing this thesis, this section was moved to an appendix for the sake of the reader. The individual patient model, physician GUI, and synthetic patient population generator were the result of previous work and were only scarcely adjusted throughout this work (e.g., to fix existing coding bugs).

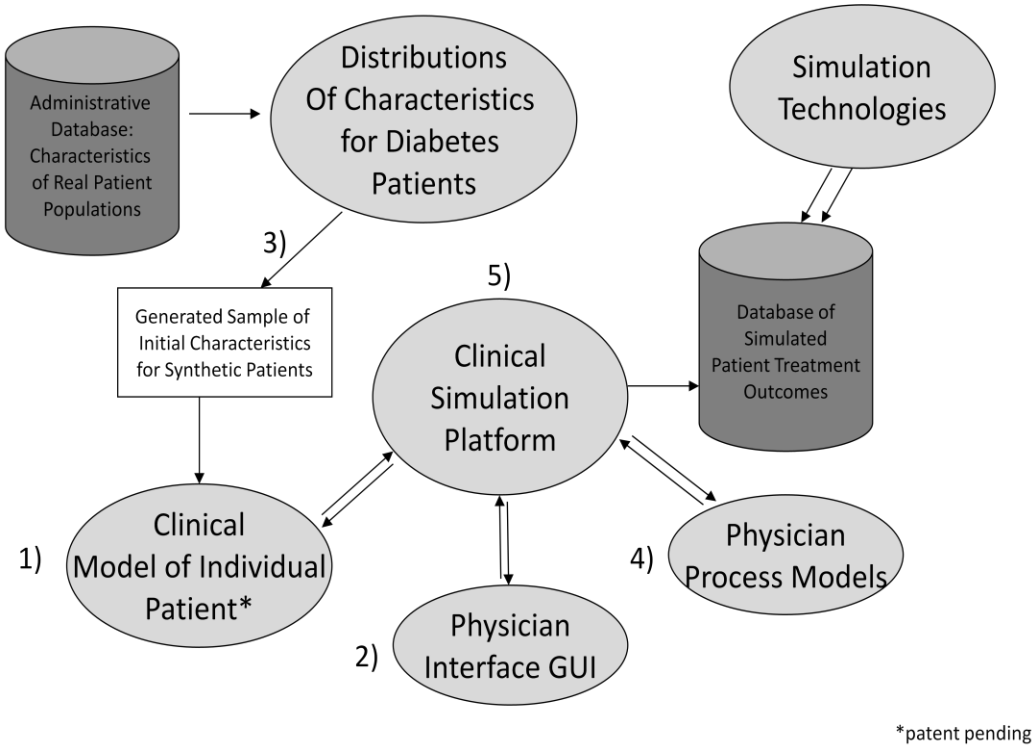
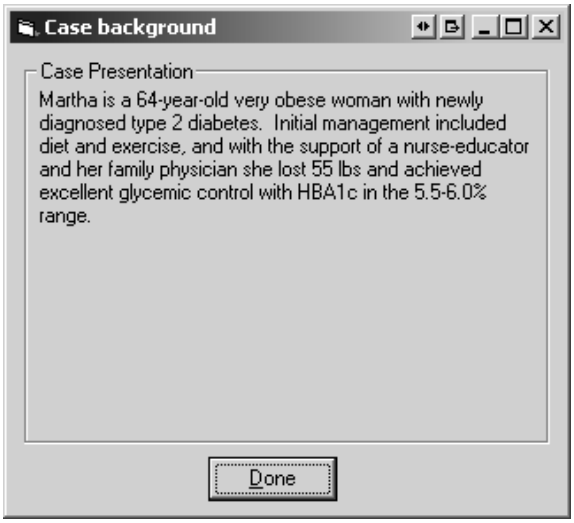


Figure 1G. An overview of the technological framework that enables the SimCare Patient Model. The SimCare model is the Clinical Model of an Individual Patient (1), and is the only piece being validated. The other components are mentioned for clarity, completeness, and credit's sake.

Real Physician Interaction with Simulated Patients

One of the intended uses of the SimCare model was to facilitate the interaction of real physicians with simulated patients. This enabled physicians to perform relevant actions on represented types of patients for the purposes of training and skill development, including patient archetypes the physician was identified to have problems treating. The GUI designed for this use granted physician access to the patient state while enabling them to order tests for current information, treat the patient using a drug formulary, and schedule future visits (Figure 2G).



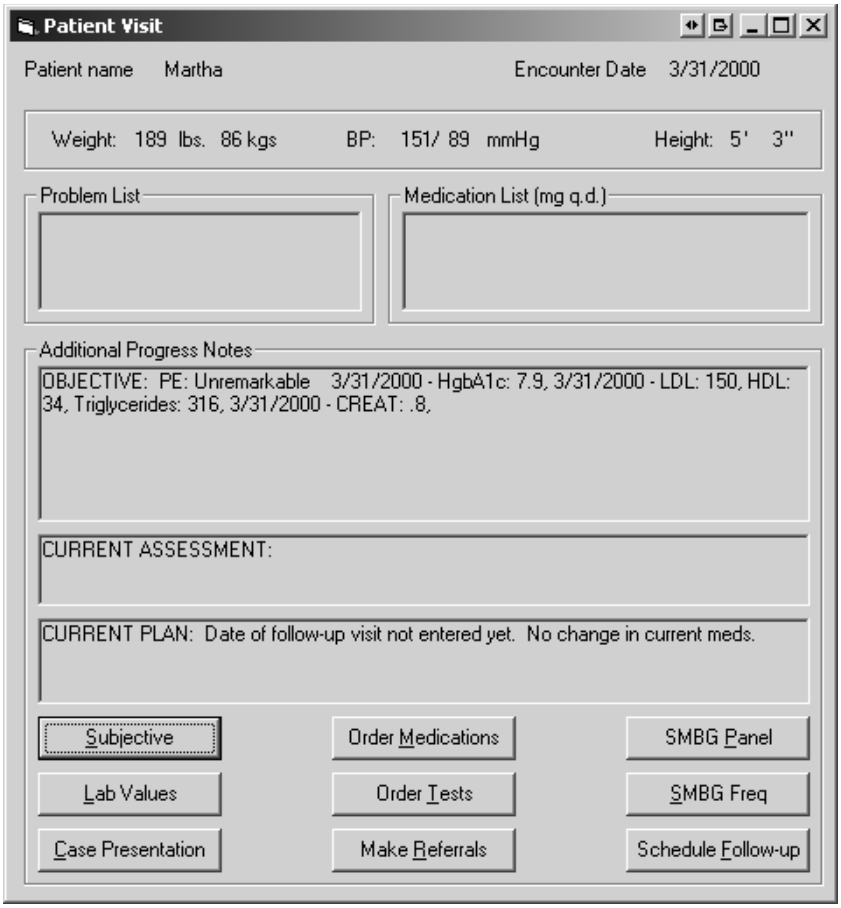


Figure 2G. Part of the physician interface is shown in the two screen shots above. The first screen presents relevant notes about the patient. The second screen enables the physician to order lab tests, medications and schedule follow-up visits.

Population Generator

The population generator is a set of correlated distributions of patient health indicators used to generate initial values for an arbitrary size of simulated patient population. The distributions were formed using samples of de-identified data from a real, clinical database across the relevant and available data attributes. Two main distributions are used. The first (Figure 3G) is modeled after the distributions of point-in-time, cross-sections of real clinical databases and is used to simulate results that represent real, clinical results. Areas of the patient space that are densely populated with patient data in the clinic are also densely populated with data in the simulated population. The second is a uniform distribution (Figure 4G) used to fully test the different treatment strategies throughout the space of possibilities. A uniform distribution is created by evenly distributing a set number of data point (representing patients) over a defined range of space. This distribution is used when areas of the patient space are too sparse to generate sufficient data for specific kinds of patients (e.g., A1c <= 11.0%, Adherence < .50).

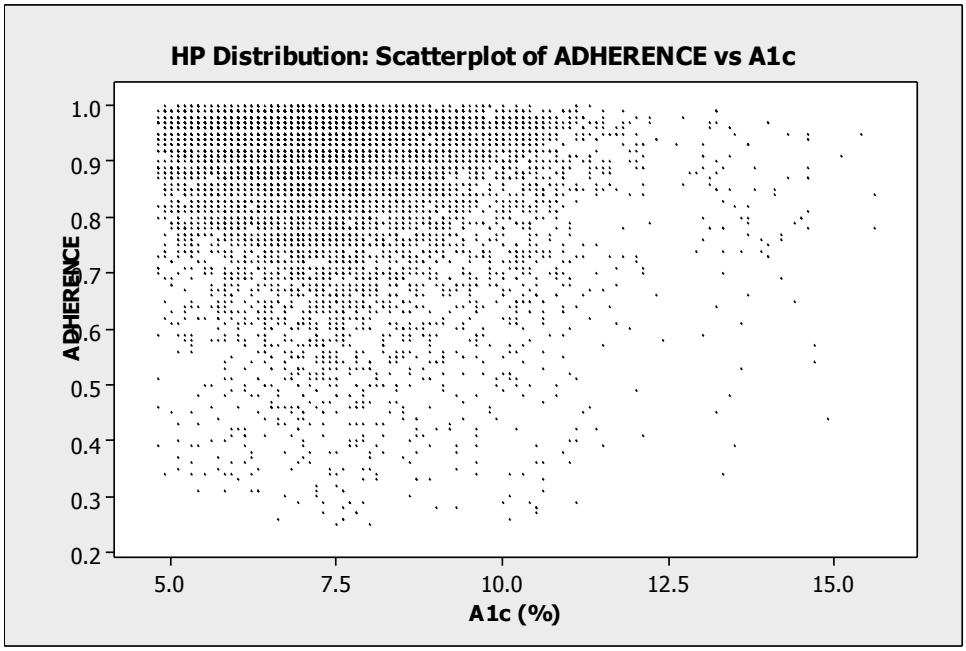


Figure 3G. A simulated population of 10,000 patients with initial A1c and Adherence values sampled from distributions modeled from real clinical data.

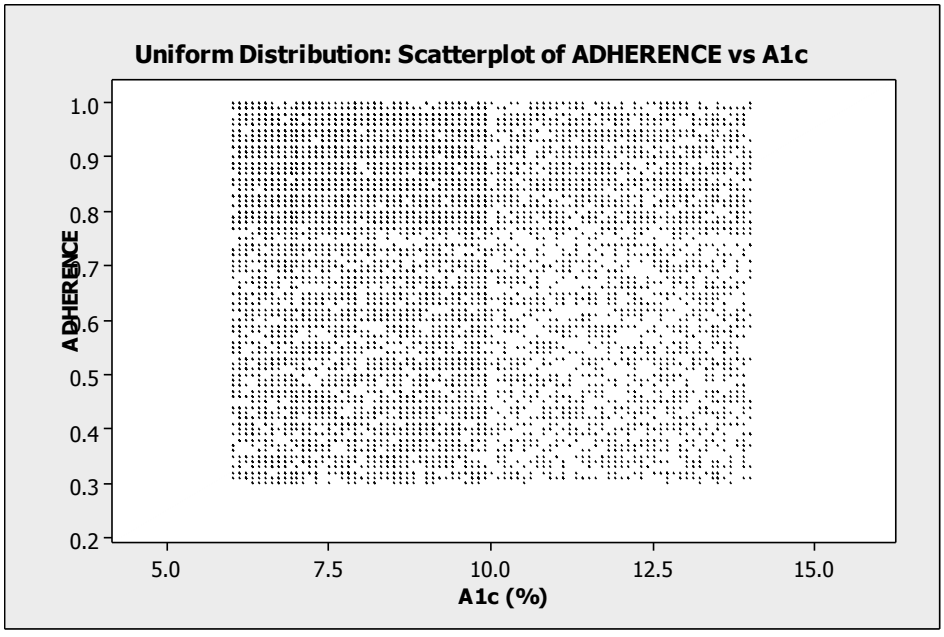


Figure 4G. A simulated population of 10,000 patients with initial A1c and Adherence values sampled from a mixture model of uniform distributions. Mixtures were based on physician descriptions of patient categories (e.g., A1c > 10%).

Physician Process Models (PPMs)

The Physician Process Model (PPM) is an implementation of a control theory formulation of a canonical clinical treatment protocol. The description of these models is included as an appendix because 1) variations of these models were used to generate experimental data for much of the thesis, and 2) they represent a significant amount of work, collaboration and contribution by the author. A deep description and background of dynamic decision making or control theory is not provided by this work. Instead, instances of the decision-making models are contrasted for the purposes of basic illustration.

Since type 2 diabetes cannot be cured but rather must be managed over time in an effort to maintain patient states, control theory lends itself in a straightforward way to model this problem. Figure 5G shows a control theory model of type 2 diabetes management. The model has three components: inverse model, patient model, and mental model.

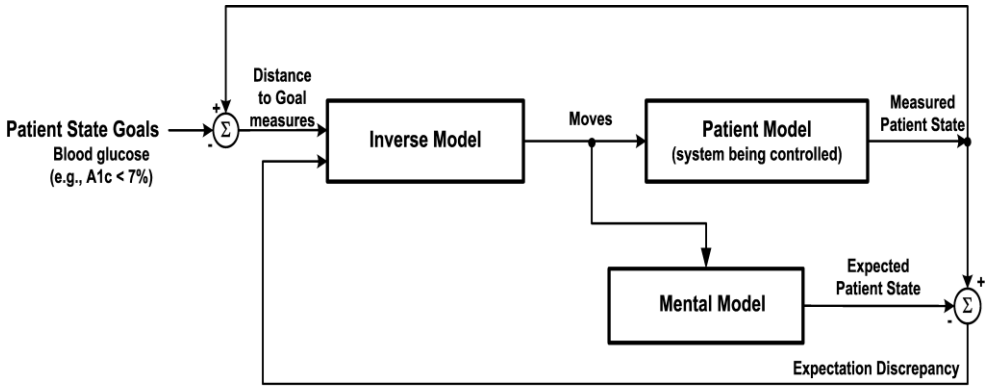


Figure 5G. Physician Process Model for managing patients with type 2 diabetes.

The *inverse model* is the component that receives patient state information and computes the actions to take in order to bring the patient to a targeted goal. Actions the model may take include medical moves (described previously as oral and insulin moves), information gathering moves and scheduling moves. Medical moves exogenously change the patient state. Information gathering moves (e.g. ordering A1c tests in advance) give the inverse model the most recent, accurate information available. Scheduling moves control the pace with which the other decisions may be made. In this way, the inverse model computes a series of dynamic decisions in real time in order to bring the patient to the goal state.

The *mental model* is the representation (set of knowledge about the control process and the process being controlled) that the inverse model maintains of the patient state and is what generates an expected state resulting from the actions of the inverse model. (The term "mental model" can be used to describe aspects of dynamic decision-making *in mente*, however the term here is used for convenience to describe a computational aspect of the decision-making model.) The mental model ranges from simple to complex, depending upon the way the inverse model generates actions. Conant and Ashby argue that every good regulator of a system must be a model of that system (Conant and Ashby 1970), but Brehmer cautions that such models are only aspects of control strategies (Brehmer 1990). Thus, mental models may range from simply computing a medical move should lower A1c to computing the future health states of the patient using dose response and time effects estimations for previous medical moves that have not yet manifested through feedback.

The *patient model* is the model of the patient that has been previously described. In this work the main aspects of the patient model that matter are tracking the A1c in response to medication moves, adherence levels and disease progression, tracking adherence levels as they vary over time and tracking all other attributes defining the patient throughout the clinical simulation.

The *expectation discrepancy* information is obtained by comparing expected resulting patient states from the mental model with the actual resulting patient states. Different control strategies use this information in different ways as described in more detail below.

Feedback Models

Models of feedback control strategies (Brehmer 1990) use current information about the patient state to solve the inverse problem (e.g., lowering the patient's A1c level to the targeted goal). The difficulty with solving this problem is that a control system cannot administer medication to the patient model without current, accurate patient state information, or else the A1c level might decrease to a dangerous, hypoglycemic level. Because each patient reacts differently to a given medication dose, the effect of the medication action in the patient state is not easy to anticipate (via heuristic or look up table). When the patient model is in such an "opaque" condition, a feedback model will have to schedule an additional encounter after enough time has passed so that accurate

patient state information can be used for the next set of computations by the inverse model.

For this work, feedback models were designed with the assumption that the time dynamics of the responses of the patient state to available control actions should be at least as fast as the opportunities to gather information and compute next actions. In this case, the dose response curves for oral medications are approximately 90 days. Thus, feedback control strategy models would not make subsequent medical moves any closer together than 90 days in order to incorporate accurate feedback into the decision making process.

Another aspect of feedback decision making is "anchoring and adjustment" (Tversky and Kahneman 1974). This part of a feedback strategy links one aspect of the dynamic patient state (the anchor) to the triggering of an action (adjustment) by the inverse model. This work uses two simple forms of anchoring and adjustment to define two different feedback control strategy models. The first form (the weak form) computes the direction the simulated patient's A1c has moved since the last encounter and uses that direction as a basis for computing actions. If the patient's A1c has stopped decreasing and is still above the stated goal, then a medical move is made. The second form (the strong form) anchors the computation only on the patient's A1c being above the goal. If the patient's A1c is above the stated goal, then an additional medical move is made.

As stated earlier, in order for feedback strategy models to operate optimally, they cannot successfully operate within a delay of feedback. Thus, both feedback models only schedule encounters with patients every 90 days (at minimum). The expectation discrepancy between the expected and actual patient state is computed to determine if the patient is responding to the medication in the expected way. An initial assumption of the inverse model is that a novel patient has a high adherence level. If a simulated patient's A1c does not lower as a result of a medical move, then the inverse model applies the heuristic that this simulated patient has low adherence. The estimate of the patient's adherence level is then re-computed to account for the discrepancy.

Feedforward Models

Models of feedforward control strategies (Brehmer 1990) use information about both the current state of the simulated patient as well as anticipated future patient states in order to compute treatment actions. This contrasting feature with the feedback model allows the feedforward model to operate within a delay of feedback period. As mentioned earlier, such a strategy model requires a more complex mental model in order to compute an estimate of the simulated patient's state before new feedback is given. An example of a more complex mental model is one having intermediate points along the dose response curve (e.g., expected results of an oral medication move at 30, 60, and 90 days). Access to this more complex mental model of the simulated patient response allows the feedforward model to generate actions more often than a feedback model.

This work has defined two feedforward control strategy models. The weak form uses a mental model that contains intermediate points along the dose response curve at days 60 and 90, thus the inverse model may schedule future encounters and make additional moves as near as 60 days after the current encounter. The strong model of this feedforward control strategy contains intermediate estimates of the dose response curve at 30, 60 and 90 days and can, therefore, schedule and compute medical moves accordingly.

Both feedforward control strategies make use of the expectation discrepancy. If the expected results of a sequence of previous moves are not met in two consecutive encounters, then the estimate of the patient's adherence level is re-computed (via heuristic) to reflect the simulated patient's low adherence. Furthermore, under the condition where the expected result has not been observed, the feedforward strong model then computes two medical moves (by titrating two separate drugs in parallel) at the same time to compensate for the lack of simulated patient progress achieved so far. Figure 6G shows an example configuration for the decision rules of four physician process models, variations of which were used to generate many of the experiments of this thesis.

- Feedback Weak
  - Getting to Goal: Make a move when A1c above goal and has stopped decreasing
  - Smooth Landing: Same as getting to goal
  - At Goal: If patient's A1c rises above 7.5%, make a move
  - Scheduling: Schedules every 90 days
- Feedback Strong
  - Getting to Goal: Make a move when A1c is above goal
  - Smooth Landing: Same as getting to goal
  - At Goal: If patient's A1c rises above 7.0%, make a move
  - Scheduling: Schedules every 90 days
- Feedforward Weak
  - Getting to Goal: Make a move when A1c is above goal
  - Smooth Landing: when A1c < 7.5 model waits to make move when A1c has stopped decreasing
  - At Goal: If patient's A1c rises above 7.5%, make a move
  - Scheduling: Schedules every 60-90 days (It schedules 90 days if the expected state from the previous visit is obtained. If not it schedules 60 days)
- Feedforward Strong
  - Getting to Goal: Make a move when A1c is above goal
  - Smooth Landing: Adds Dose Response curves to achieve smooth landing
  - At Goal: If patient's A1c rises above 7.0%, make a move
  - Scheduling: Schedules every 30-60 days (It schedules 60 days if the expected state from the previous visit is obtained. If not it schedules 30 days)

Figure 6G. An example configuration of the decision rules of four Physician Process Models (PPMs). Getting to Goal is the decision rule to trigger additional medication moves from an initial patient state. The Smooth Landing rule describes how the model attempts to prevent the simulated patient from entering a hypoglycemic state (over-medicating the patient). The At Goal rule defines how the PPM maintains the patient state after the targeted goal is reached. Scheduling rules are also defined.

Clinical Simulator

The clinical simulator controls the interface between the PPM and the patient population and is designed so experiments can be run under varying simulated clinical and decision making conditions. The parameters for the specific PPM to be tested are selected, and various patient subgroup characteristics are selected as well. Parameters are set to differentiate experimental clinical environments, such as drug formularies and dose increments, adherence treatment options, scheduling options, test ordering options, number of treatment encounters to simulate and evidence-based goals (Figure 7G). A drug formulary is a listing and ordering of drugs that are available for use by the

Physician Process Model. An example of a formulary with three available drugs is Metformin, Glipizide, and Rosiglitazone. Formularies to treat lipids and blood pressure are defined separately in a similar manner. (As mentioned in the opening of this appendix, this description is not a full, detailed description of the PPM or of the related pieces of the models that enable the simulations to be run. As can be seen in the figure below, many parameters and options exist to generate – and differentiate – clinical simulation environments.)

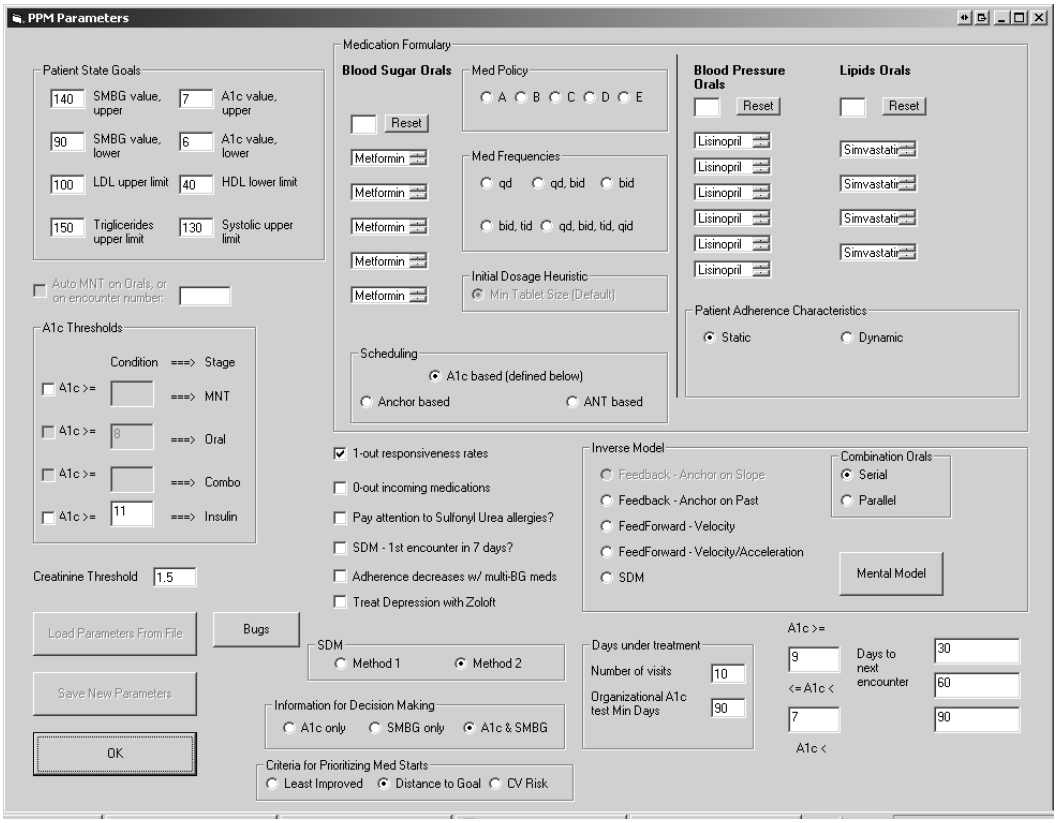


Figure 7G. A screen shot of the clinical simulator GUI. The parameters establish the type of PPM, some patient characteristics not already in the simulated patient file, the evidence-based goals and other clinical parameters for treatment of the patient population. Output are treatment records.

After the PPM parameter graphical user interface (GUI) is set and the file of synthetic patients is loaded, the clinical encounter simulation is run. The simulation works by the PPM treating a patient over the treatment time specified (number of encounters) and then moving on to the next synthetic patient in the file. The full results of these treatments are output into text files and are loaded into a database for analysis. These treatment records are used to analyze both the decisions that were made along the treatment path as well as the resulting patient outcomes.

Figure 8G depicts an A1c treatment path for a simulated patient in the SimCare model. This patient had an initial A1c = 10.0% and was treated for 12 years of simulated time. Clinical encounters are denoted by circles. In this case, the patient was first given oral medications to attempt to control their blood sugar. The evidence-based goal was not met and the oral medication formulary had been exhausted, so then insulin was administered to bring the patient to the A1c goal. Insulin was then routinely adjusted to manage the patient's blood sugar and keep them at goal.

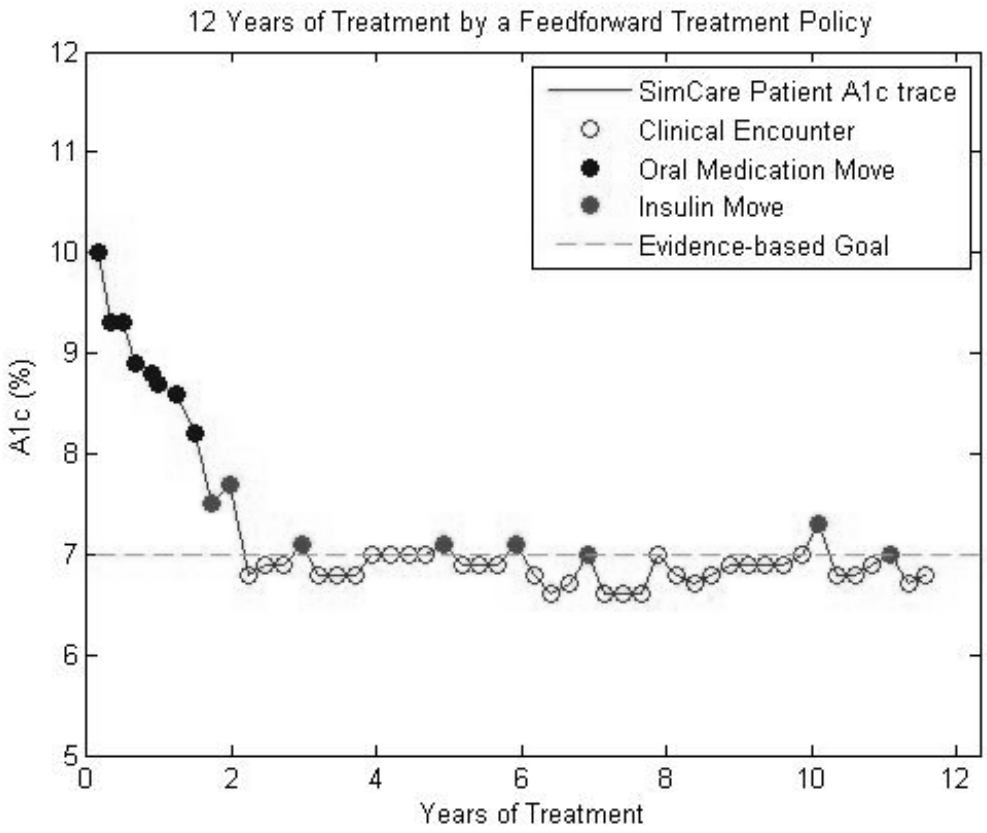


Figure 8G. 12 years of treatment of on an individual SimCare patient. Each circle represents a clinical encounter and generated patient state. Dark circles are encounters where oral medications are increased; light circles, insulin is increased; hollow circles, medications are not adjusted.