

## **APPLYING A FRAMEWORK FOR HEALTHCARE INCENTIVES SIMULATION**

Joseph P. Bigus  
Ching-Hua Chen-Ritzo  
Keith Hermiz  
Gerald Tesauro

Robert Sorrentino

IBM T.J. Watson Research Center  
1101 Kitchawan Road  
Yorktown Heights, NY 10598, USA

IBM T.J. Watson Research Center  
19 Skyline Drive  
Hawthorne, NY 10532, USA

### **ABSTRACT**

At WinterSim 2011, we originally proposed an agent-based framework for healthcare simulations, enabling flexible integration of multiple simulation models, including models of disease progression, effects of provider interventions, and provider behavior models that are responsive to contractual incentives. In this paper, we report results using our proposed framework to integrate two examples of provider behavior models, two examples of disease models, and four examples of payment models. We explore multiple combinations of these models and simulate the impact that alternative payment models may have on health and financial outcomes. These examples test the robustness of the simulation framework, and illustrate the value of such simulations to the policy makers who design incentives to improve cost and health outcomes, and to providers who wish to evaluate the financial impact of proposed incentives on their practice.

### **1 INTRODUCTION**

The healthcare industry is a complex system involving payers, providers, and patients. This system is under tremendous pressure for change as consumers, employers, and government health policymakers try to control rapidly increasing costs while simultaneously maintaining or improving health outcomes. A wide variety of major structural changes have been proposed, potentially impacting payer revenue and costs, provider revenue and treatment plans, and patient costs and health outcomes. Using the flexible simulation framework and Eclipse-based development tools introduced in Bigus, Chen-Ritzo, and Sorrentino (2011), we have built a set of parameterized health system models that can be configured by executive-level decision-makers to ask and answer what-if questions in realistic scenarios pertinent to their business.

In this paper, we use the framework of Bigus, Chen-Ritzo, and Sorrentino (2011) to implement two examples of provider behavior models, two examples of disease models, and four examples of payment models. We use these models to simulate the impact that alternative payment models may have on health and financial outcomes. Since our simulations are not based on real data, our objective is to use these examples to test the robustness of the simulation framework and illustrate the potential value of such simulations to health policy-makers and health providers. Health policy-makers can use such simulations to aid the design of provider incentives to improve population cost and health outcomes. Meanwhile, providers may use such simulations to evaluate the financial impact of proposed incentives on their practices.

A depiction of the basic interaction cycle between components in our simulator can be seen in Figure 1. A population of Patients interacts with a healthcare Provider over a sequence of several decision cycles, typically of one year duration. Based on observation of the patient types and health states, the Provider will select from among various options for treatment of each Patient. The selected treatments will result in claims being submitted to a Payer, which will then issue payment based on the payment model being used. Note that the payment model may take into account health states and outcomes, along with the actual

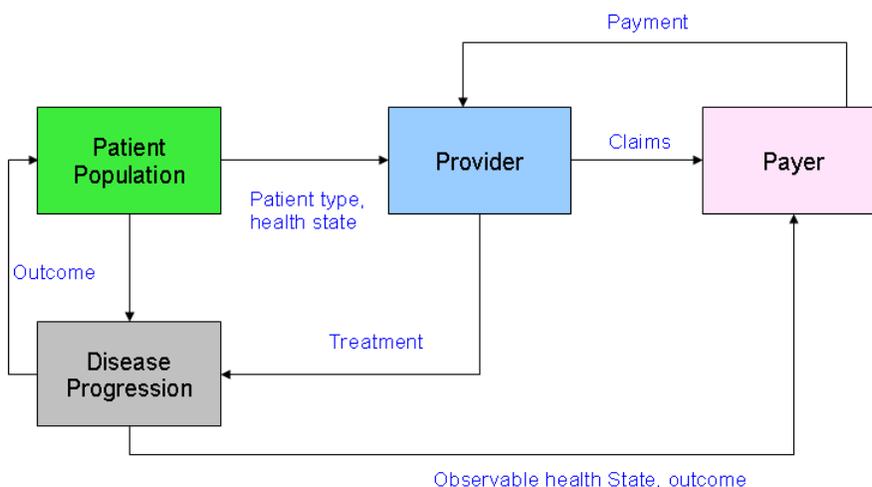


Figure 1: Basic interaction loop in simulated healthcare scenarios.

treatment procedures, in computing amounts paid. Outcomes of treatments, along with recalculated health states for the next decision cycle, are computed using a Disease Progression model, based on the current health states and the selected treatments.

The key assumption that we make in determining how payment models affect financial and health outcomes is that the Provider selects treatments on the basis of optimizing a utility function that is specified by the payment model and the Provider’s profit margin on the various treatment options. As such, the Provider’s objective in our simulations is solely to maximize profitability, but non-financial objectives such as concern for patient welfare could also be included in the utility function definition. We compare two different optimization methodologies: a one-step “myopic” optimization, based on a known model of immediate outcomes of treatments, and a Markov Decision Process (MDP) formulation that aims to optimize cumulative utility over multiple time steps.

The rest of the paper is organized as follows: In Section 2, we motivate our work by presenting and describing screen shots of the main simulation outputs that policymakers and healthcare providers can use. In Section 3 we briefly describe the two versions of diabetes models that we utilize in our simulations. This is followed by a section that describes the available provider interventions for each of the diabetes models. The different forms of provider incentives used in our simulations are presented in Section 5. Sections 6 and 7 present the algorithms used for simulating provider behavior, along with the corresponding simulation outputs. Finally, we conclude this paper with a summary discussion.

## 2 MEASURING THE IMPACT OF INCENTIVES ON HEALTH AND COST OUTCOMES

In this section, we describe a decision-support tool for use by healthcare executives, chief medical officers, chief financial officers, policy strategists, and government agencies to aid in understanding the impact of proposed changes to the standard fee-for-service payment model.

Figure 2 shows a screen shot of a web-based user interface or dashboard to specify scenarios for simulation and to explore the financial and health outcome results. The dashboard comprises four major areas: on the left hand side, a set of panels containing widgets or user-interface controls for configuring the scenario; on the right hand side, a set of tabs containing graphs or charts showing results from the simulation; on the top, a set of buttons for controlling the simulation run; and on the bottom, a status field that shows progress and information regarding the current simulation.

The configuration controls include five panels, business models, patients, providers, payers, and simulation. The business models section allows the user to specify two major parameters in the simulation,

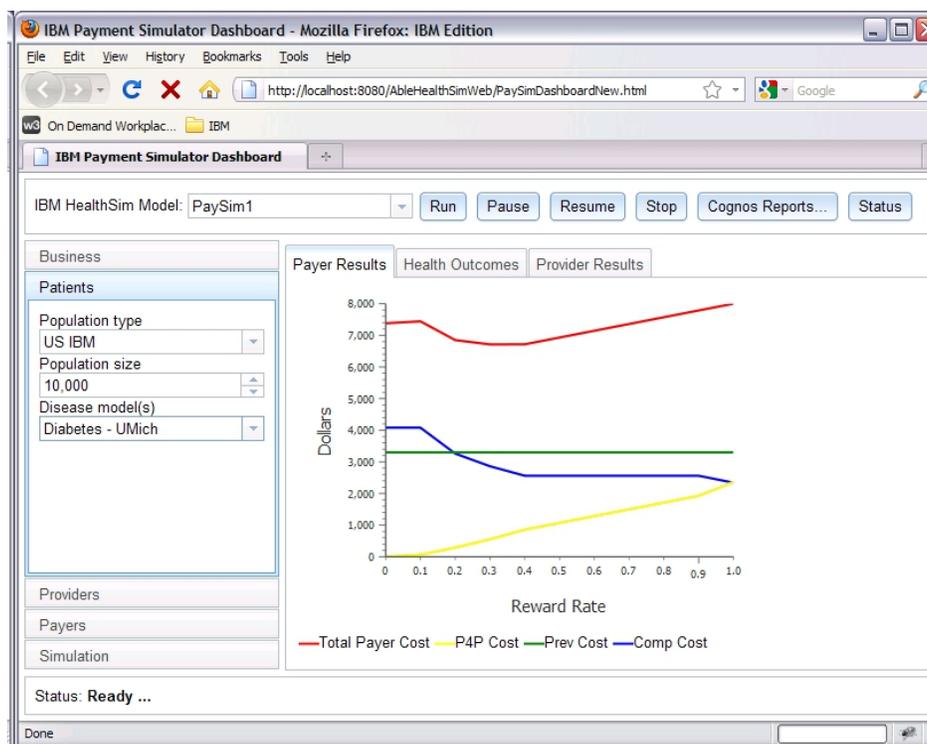


Figure 2: Executive dashboard for payment incentives analysis.

the type of Provider business organization and the clinical model used by primary care physicians. The Patient parameters, as illustrated, include the patient population type (representative of a particular country or locale, employer, or specific disease group), the number of patients in the population, and the specific disease models to be used for disease progressions and treatment impacts on future health states. The Provider section specifies behavior models used for learning decision-making models and policies, and the number of providers. For Payers the single parameter specifies the reimbursement model to be used for the selected Providers. Finally, the Simulation parameters include the forecast period (number of steps, typically years) and the number of simulation runs to perform.

The configuration parameters are then mapped to a single simulation model instance, which is loaded and run on the web server (cloud computing) environment. When the set of runs are complete, summary results charts are displayed in the right hand tabs. In the example shown in Figure 2, these include a set of Payer Results, the patient population Health Outcomes, and the resulting Provider Results. Key performance indicators for Payers would include revenue vs. costs (profits), any risk-sharing payments to providers, the medical loss ratio, etc. For Providers, the results would include revenue vs. costs as well as outcome based measures such as patient panel biomarkers (e.g., average Hemoglobin A1C). Patient Health Outcomes would present summary data on the patient population health states and levels of complications (hospitalization, non-elective emergency treatments, etc.). The Cognos Reports button allows custom business intelligence reports to be run for further analysis.

The purpose of the dashboard interface is to bring the locus of simulation control to the end users, instead of requiring specialized programming for each variant of the model being used in the simulation. In this way, a health policy leader could use our platform to forecast the clinical and financial outcomes of managing one or more segments of the eligible population (e.g., all the diabetics in a given geographic region), by specifying appropriate characteristics of the healthcare ecosystem. Our objective is to provide a degree of usability to our platform that is seldom seen in healthcare simulation.

### 3 DIABETES MODEL

We consider two disease models in our simulations, both of which are Markov models of diabetes. The first of these models was created at the University of Michigan (Barhak, Isaman, Ye, and Lee 2010) and was calibrated against real-world clinical studies. The second model is a much simpler model created at IBM that has not been calibrated against any real-world data. While the calibrated model is more realistic, it is also more complex and its dynamics are more difficult to control using a few input parameters. The Michigan model additionally includes continuous state variables such as blood sugar levels, which make it difficult to provide a formal MDP specification in terms of enumerated states and state transitions. Conversely, all state variables are discrete in the IBM model, allowing easy application of formal MDP solution techniques.

An illustration of the simplified IBM diabetes model is provided in Figure 3. This model contains six states representing various stages of diabetic retinopathy, ranging from “No Retinopathy” to “Blindness.” We consider the state “Macular Edema Entered” to be a complication state. Whenever a patient enters this disease state, it triggers a series of mandatory provider interventions, followed by entering the “History of Macular Edema” state with probability 1.0. The reader is referred to Barhak, Isaman, Ye, and Lee (2010)

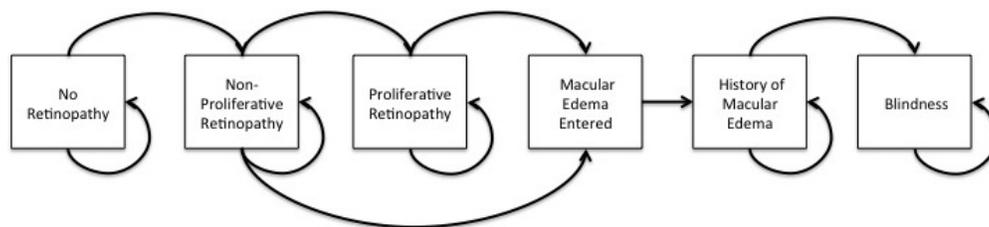


Figure 3: Simplistic model of retinopathy pathway for diabetes.

for more details on the Michigan diabetes model, which includes multiple disease pathways, as is typical in diabetes modeling research. The reasons that we chose to experiment with two different disease models is so that we could demonstrate that our simulation framework was flexible enough to accommodate different instantiations of a disease model. The fact that one of these models was extremely simple and easy to control made it much easier for us to develop and debug our code, before we attempted to run simulations with the more complex and realistic disease model.

### 4 PROVIDER INTERVENTIONS

In this section, we describe the elective and non-elective provider intervention models used in our simulations. We first describe the modeling of elective interventions. In general, the available elective interventions match interventions that the disease model was designed to support. For example, the Michigan model supports three interventions (diet and exercise education, oral medication, and insulin), corresponding to treatments that were used in the real-world clinical study that the disease model was based on.

Meanwhile, for the simple diabetes model, we invented four effective treatment options, which represented various combinations of two types of treatments, each of which could be implemented at two different levels of intensity. The two types of treatment are medication management (MM), and self-management training (SMT). We assume that each treatment is applied at either of two levels (i.e., 1 = high and 0 = low) of intensity, simultaneously. Therefore, the four treatment combinations are: MM-0/SMT-0, MM-1/SMT-0, MM-0/SMT-1, MM-1/SMT-1.

For both diabetes models, we assume that all treatments are feasible options in all patient states. The disease-state transition probabilities are dependent on the selected treatment. Since the Michigan model

does not specify the reimbursable procedures that are associated with each treatment, we implemented our own set of procedures. As described in Bigus, Chen-Ritzo, and Sorrentino (2011), every treatment may be decomposed into a sequence of encounters (or visits) with a provider, and each encounter may be decomposed into a collection of actions, i.e., billable procedures that can be claimed from the payer. We specify both a reimbursement rate and a provider cost for each procedure.

With respect to non-elective interventions, we assume that these interventions only occur in the disease state ‘Macular Edema Entered,’ which exists in both disease models. The patient enters this state with some probability, but with probability 1.0, a non-elective (mandatory) treatment comprising of a series of laser eye treatments will be implemented by an ophthalmologist.

## **5 INCENTIVE MODELS**

We consider three types of provider incentives. The first is a fee-for-service (FFS) incentive, where providers are reimbursed a fixed amount from the payer for each procedure or service that they perform. Hence, provider revenue is directly proportional to the number of procedures performed, and the profitability per procedure. Under this scheme, there is no (financial) incentive for the provider to limit the number of procedures performed. However, in practice, aside from the provider’s own code of ethics, providers face capacity constraints, and there exist laws and other policies that may bound the number of procedures that can be delivered to a single patient.

The second type of incentive is the global capitation (GC) incentive. In this case, providers are paid a flat rate for each patient each month, regardless of the number of procedures or services that may be delivered to the patient. This rate is called the capitation rate, and may vary by the patient’s general state of health. Under this scheme, there is no (financial) incentive for the provider to provide any services to the patient. However, again, there are practical reasons why providers reimbursed under this scheme still provide care to their patients.

The third type of incentive, called the ‘shared risk’ incentive, is applied as a variation to the previous two incentives. When shared risk incentives are applied, the provider is liable for a portion of the losses or savings achieved, relative to a pre-defined threshold. We refer to this threshold as the ‘reference cost’. In a fee-for-service with shared risk incentive scheme, the provider is credited in a fee-for-service fashion, with the exception that a portion of any positive difference between the reference cost and the fee-for-service credit (i.e., the savings) is added to the provider’s fee-for-service reimbursement as a bonus. The fraction of the savings that the provider is eligible for is referred to as the ‘reward rate’. At the same time, a portion of any negative difference between the reference cost and the fee-for-service credit (i.e., the excess spending) is deducted from the provider’s fee-for-service reimbursement. The fraction of the excess spending that the provider is liable for is referred to as the ‘penalty rate’.

Since the the shared risk incentive may be included with the fee-for-service or global capitation incentive, we present simulation results for provider responses under four incentive scenarios: 1) fee-for-service; 2) fee-for-service with shared risk; 3) global capitation; 4) global capitation with shared risk.

## **6 THE MYOPIC PROVIDER**

In this section, we present simulation results for scenarios in the simple IBM disease model using a “myopic” provider, whose choice of treatment maximizes utility in the current decision period only. As detailed in the Appendix, exact computation of the utility-maximizing treatments is formulated as a mixed-integer linear program, and is implemented using CPLEX. This formulation operates over the entire population of patients to be treated in the current decision period, and takes into account any capacity constraints, such as the maximum number of appointments that can be scheduled per period.

The IBM diabetes model, as mentioned previously, assumes four possible treatments that vary in treatment intensity. As detailed below, higher intensity treatments yield greater profitability for the provider, as well as generally slower progression through the disease states. Hence, in the absence of capacity constraints, a

provider with fee-for-service incentives will always select the maximum intensity treatment for every patient. However, in certain states of the model, higher intensity treatments do not improve patient health outcomes. For example, once a patient reaches the Macular Edema Entered state, all further disease progression is independent of selected treatment. This allows the possibility of reducing the total cost of care without harming patient outcomes, by changing the provider incentives so that less expensive treatments are selected in these states.

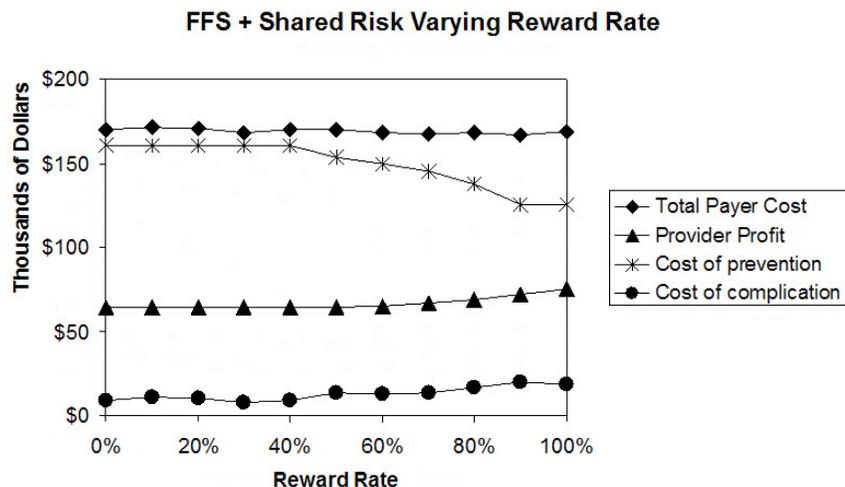


Figure 4: Results for FFS+Shared Risk vs. reward rate, using myopic optimization of the provider’s treatment policy.

The fees for preventive care treatments are largely driven by the medication management intensity. The fee for the most aggressive treatment is roughly \$1,000 while the least is approximately half. All the fee-for-service simulations exhibited were based on a reference cost of \$7,500 per patient per year and assumed provider cost of service equal to 60% of fees associated with the chosen treatments. The reference cost will support half of the patient population at the highest intensity of medication management provided there were no complications. The following results represent an average over 40 simulations of a single provider with a panel of ten patients progressed over 15 years.

The addition of a shared reward component to the fee-for-service intervention causes the primary care provider to trade off the profit from treatments with higher intensity against the potential share in a savings pool. The provider is eligible for a reward if the total cost of patient treatment across her panel, both elective and non-elective, comes in below the reference cost. The trade off is not attractive for very low shared reward rates, as seen in Figure 4. When the reward rate is below the provider profitability (40% of fees) her incentive is to apply the maximum intensity intervention to each patient. Once the reward rate reaches the providers profitability she begins to reduce the intensity level of treatment on selective patients in favor of sharing in the pooled savings. As she continues to reduce the level of intensity of preventive care across her patient panel complications, and the cost of non-elective interventions, rise. These costs reduce the pool but the provider’s profit continues to rise as the reward rate is increased. In the end, she has managed the intensity of treatments so that half of her patients are at the highest level and half at the lower. Note that the total payer cost is nearly constant.

Figure 5 shows the results for a penalty based fee-for-service incentive model. In this scenario the provider shares the cost of elective treatments that exceed the reference cost. The provider once again continues to treat her entire patient panel at maximum intensity at the lower penalty rates. However, her profitability is being impacted immediately. If she treats her patient panel in a way that results in the cost

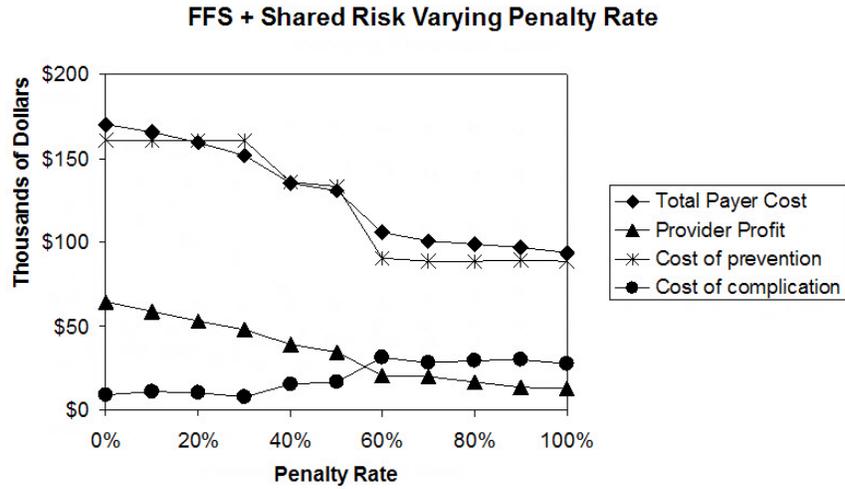


Figure 5: Results for FFS+Shared Risk vs. penalty rate.

exceeding the reference, she recoups her 40% profit but now is liable for a penalty as a percent of the overrun cost. As the penalty rate exceeds the profitability rate she loses money on those treatments that exceed the reference cost and she changes her strategy. She aggressively cuts back on treatment intensity across her patient panel until she is treating each patient at the minimal level. The number of complications and their associated cost rise and finally level off.

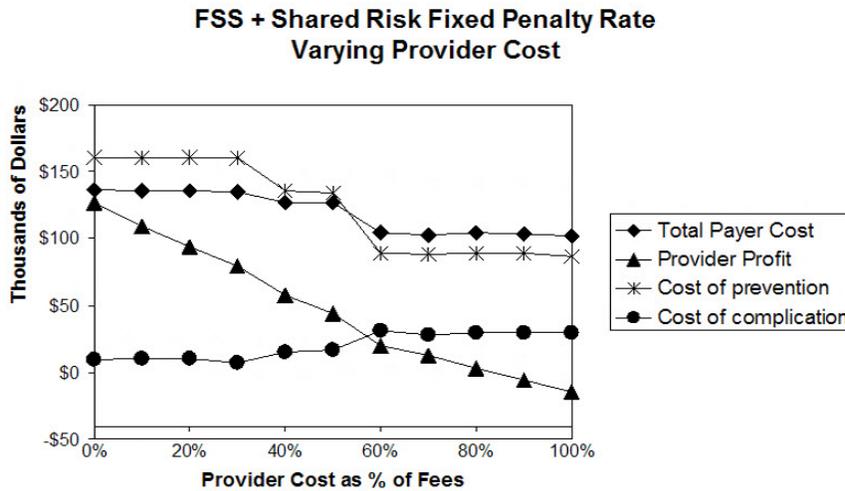


Figure 6: Results for FFS+Shared Risk vs. provider cost.

We can further examine the role that the provider’s profitability plays in the selected treatment policy. Figure 6 shows the prior fee-for service with penalty incentive model with the penalty rate fixed at 60%. In this case we vary the provider’s cost as a percentage of elective fees. A provider with very low cost of service will be largely immune to incentives, whereas a provider with very high cost of service will be highly compliant. However, that compliance saturates when all patients are being treated at the minimum level of intensity. The incentive model can achieve compliance, cost savings, and adequate care without

adversely affecting the financial health of the provider only if the provider's costs are not at the extreme highs or lows.

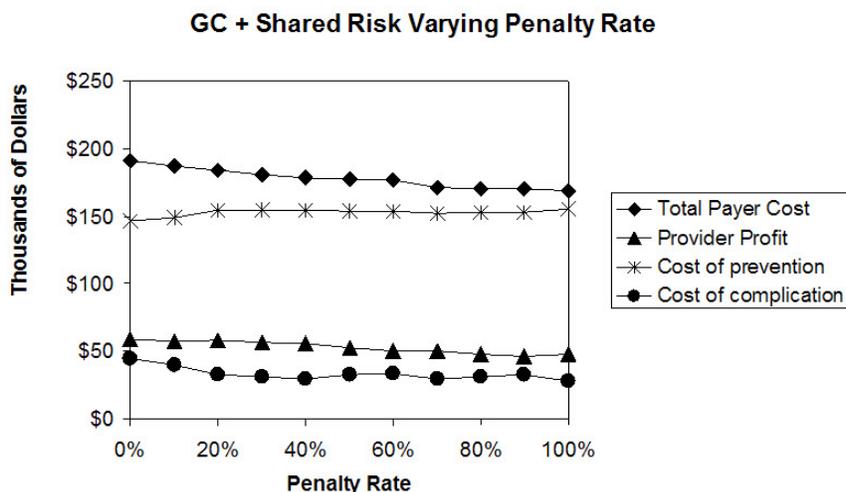


Figure 7: Results for GC+Shared Risk vs. penalty rate.

A carefully crafted global capitation incentive model can balance the interest of patient, provider and payer. Beyond establishing a reference cost and a rate for shared risk, the payer also has to establish appropriate flat rates for care for patients in various health states. These bundled payments are significant levers in encouraging provider behavior and need to consider downstream effects on patients health. Figure 7 shows the results of simulations under a global capitation incentive model with penalty. The reference cost was set to \$10,000 for these runs. As the penalty rate is increased we see a smooth change in the provider's treatment policy which results in minimally increasing prevention costs while decreasing complication costs. The overall cost to the payer decreases and the impact on provider profitability appears to be small.

## 7 REINFORCEMENT LEARNING PROVIDER

We now present a provider decision model based on an MDP formulation of treatment selection, which aims to optimize long-range cumulative utility, wherein future utility may be progressively discounted by a discount parameter  $\gamma \leq 1$ . This allows leveraging a vast literature covering a wide variety of techniques for exactly or approximately solving for optimal policies in MDPs. The MDP solution technique used here is based on Reinforcement Learning (RL) (Sutton and Barto 1998), a trial-and-error technique known to effectively optimize policies based on extensive simulation trials. The main advantages of the RL approach are that an explicit formal specification of the MDP is not needed, and it effectively tackles large-scale MDPs via use of linear or nonlinear function approximation.

For the IBM diabetes model, we specifically use Q-Learning (Watkins 1989), which learns a table of values  $Q(s, a)$  estimating the long-range value of initially performing action  $a$  in state  $s$ , and thereafter using the optimal policy. Q-Learning is a low complexity, step-by-step learning method that is guaranteed to find the exact optimal policy given sufficient "exploration" of all available state-action pairs. Since there are six possible patient health states in the IBM model, and four available treatments, a table of 24 cells can represent the expected value of treating an individual patient. The resulting single-patient treatment policy may not coincide exactly with the optimal joint policy over a population of patients, in cases where there are constraints on the joint actions (e.g., due to limited capacity for patient visits) or non-trivial joint rewards such as aggregate bonus payments. However, such exact coincidence can be rigorously shown to occur in the Shared Risk incentive model, for the special case in which the reward rate equals the penalty

rate. In this case, the aggregate bonus payments can be partitioned into well-defined per-patient bonus payments, which are thus exactly optimized in the single-patient MDP formulation. This is the scenario that we study below.

Figure 8 illustrates results using a treatment policy trained by Q-Learning with discount parameter  $\gamma$  set to zero, so that the learned policy matches the myopic policy of Section 6. We utilize a fee-for-service plus Shared Risk payment model, where we vary both the reward rate and penalty rate, keeping them equal in each experiment. We use the same configuration of 10 patients, each beginning in the No-Retinopathy state, and being treated annually for 15 years, that was described previously. The provider’s profit margin is 40% of revenue for all available treatments. At a given value of reward rate, we first train the Q-table until converged; this usually takes at most a few hundred simulation runs. We then benchmark the resulting financial metrics over 500 additional runs. Similar to our earlier experiments, we see that a low reward rate results in the provider treating all patients with the most expensive treatment, yielding the lowest complications cost. When the reward rate exceeds the profit margin, the provider prefers to switch to the least expensive treatments, except in the Proliferative and Non-Proliferative states. As these states risk an immediate complication on the next time step, the provider prefers an increased level of treatment to reduce the immediate complication risk. The best contract from the payer perspective has reward rate and penalty rate set to 50%; this reduces the total cost of care by more than 20%. However, there are more complications and the provider profit is reduced when compared with a straight fee-for-service contract.

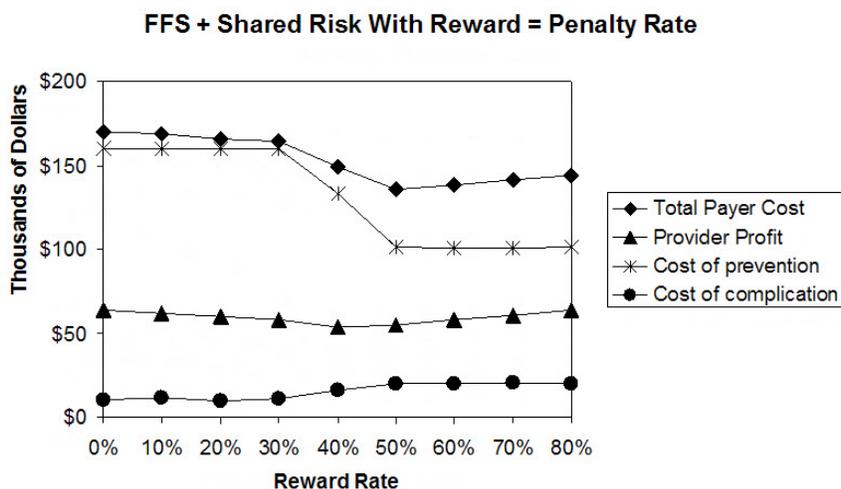


Figure 8: Results for FFS+Shared Risk vs. reward rate, using Q-Learning with  $\gamma = 0$  to optimize the provider’s treatment policy.

Table 1: Illustration of differences in results using myopic vs. lookahead-based optimization of provider treatment policies. rewardRate = penaltyRate = 0.5, provider profit margin = 40%.

Metric (K\$)	Myopic ( $\gamma = 0$ )	Lookahead ( $\gamma = 0.5$ )	Lookahead Improvement
Total Cost of Care	135.6	133.4	1.6%
Provider Profit	54.9	56.3	2.6%
Complications Cost	19.9	17.5	12.1%

Having matched the myopic policy with  $\gamma = 0$ , we then examined how the Q-Learning policy changes when we enable lookahead effects via non-zero  $\gamma$ . There is no change at reward rates less than 0.4, but above that level, we see increased provider profits and reduced complications. The lookahead-based policy

switches in the No-Retinopathy state from MM-0/SMT-0 to MM-0/SMT-1. This shows that Q-Learning is picking up the longer-term risks of complications in this state. The magnitude of improvement over a myopic policy can be appreciable, as seen in Table 1. For  $\gamma = 0.5$ , complications are reduced by over 12%, while total cost of care drops by nearly 2% and provider profits increase by nearly 3%.

## 8 SUMMARY

In this paper, we presented several examples illustrating how our healthcare simulation framework can facilitate interaction among multiple types of simulated components of a multi-agent healthcare system, including payment models, provider behavior models, and disease progression models. For an artificial model of diabetes, we also gave several examples showing how health and financial outcomes can be influenced by specific terms and parameter settings of Shared Risk payment models, and we discussed how this type of analysis can be useful to decision-making executives in the healthcare industry.

The Michigan diabetes model turned out not to be useful for purposes of our study, as all three of its available treatments yield essentially equivalent health outcomes. This leads to a trivial and unrealistic optimization of selected treatment based solely on cost. It appears upon further examination of the model's details that it presumes a specific Markov model of treatment intensification as the patient's disease progresses, and was not designed to accurately predict outcomes of arbitrary treatment policies. This realization has led us to search for other published diabetes models that include a specific focus on predicting differential health outcomes arising from different treatment alternatives. One such model that appears promising in this regard is the UKPDS Outcomes model developed at the University of Oxford (Clarke et al. 2004).

In addition to selecting and implementing a more appropriate diabetes model to simulate within our framework, our ongoing/future work also focuses on extending the provider optimization techniques to include function approximation and Approximate Dynamic Programming (Powell 2007) techniques that may be required by more complex models with continuous state variables. Also, as many Shared Risk contracts utilize a multi-year time frame for calculation of bonus payments, we are also working on requisite modifications to our MDP formulations that can accommodate such delayed reimbursement schemes.

## APPENDIX

In this section we present the formulation of the myopic provider's decision model. A primary care provider (PCP) cares for a panel of patients. We assume that the type of care provided by the PCP is preventive. If complications arise, we assume that they are treated by a non-PCP. The reimbursements paid to the PCP are referred to as the 'cost of prevention'. Meanwhile, the medical costs incurred with non-PCPs are referred to as the 'cost of complications'. The total medical cost incurred by a patient is referred to as the patient's 'system cost'.

The PCP is assumed to be a profit maximizer. She is presented with alternative reimbursement contracts by the payer and her decision problem is to select interventions for her patients in such a way that she maximizes her profit. We assume a single period decision problem in which the PCP selects an intervention policy at the start of the decision horizon.

The PCP's panel comprises  $n$  patients, which may be partitioned into  $M$  subsets, based on their health status. Let  $n_m$  be the number of patients in health status  $m$ , where  $m = 1, 2, \dots, M$ . We define an intervention very generally to comprise any combination of clinical tests, procedures and/or consultations. Suppose that there are  $K$  mutually exclusive interventions.  $x_{km}$  is the fraction of patients in health state  $m$  that will receive the  $k^{\text{th}}$  intervention. We assume that each patient must be assigned exactly one intervention. This implies that  $\sum_k x_{km} = 1 \forall m$ . Let  $\mathbf{x}_m = [x_{1m}, x_{2m}, \dots, x_{Km}]$  be the PCP's intervention for patients in health state  $m$ . We define the matrix  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]^T$  represents the PCP's *intervention policy*.

Each intervention  $k$  is associated with a requisite number of office visits. Let  $d_k$  be the number of visits required per year for intervention  $k$ . We assume that  $d_k$  is independent of the patient's health state and that a visit for one intervention occurs separately from a visit for another intervention. We consider

the PCP's time to be the primary resource constraint, so that the total demand for appointments each year, given by  $\sum_k \sum_m n_m d_k x_{km}$ , must not exceed  $A$ , the maximum available appointment slots per year.

The cost (to the payer) of prevention for all patients, given interventions  $\mathbf{X}$ , is  $v(\mathbf{X})$ . The cost of prevention will depend on the reimbursement model being used, and will be defined later. The cost of complications for a patient in health state  $m$ , given intervention  $k$ , is given by  $h_{km}$ . The probability of complications occurring in health state  $m$  with intervention  $k$  is  $p_{km}$ . The expected system cost for all patients is the sum of preventive and complications cost, and is given by  $Q(\mathbf{X}) = \sum_m [v(\mathbf{X}) + \sum_k n_m p_{km} h_{km} x_{km}]$ .

We represent the PCP's expected reimbursement from the payer by  $R(\mathbf{X})$ . In this section we present three types of reimbursement models: fee-for-service, global capitation, and shared risk. In fee-for-service, a fixed fee is paid for each procedure, test and consultation. Each intervention may comprise multiple encounters with the PCP, each comprising a collection of such procedures, tests and consultations. Let  $f_k$  be the total fees that are reimbursed to the PCP when implementing intervention  $k$ . Given the PCP's intervention policy,  $\mathbf{X}$ , the PCP's *expected* reimbursement under a fee-for-service model,  $R_1(\mathbf{X}) = \sum_k f_k \sum_m x_{km} n_m$ . In any fee-for-service setting, we equate the cost of prevention,  $v_m(x_m)$ , with the total fee-for-service reimbursements that the PCP receives. Therefore,  $v(\mathbf{X}) = R_1(\mathbf{X})$ .

Under a capitation contract, the PCP receives a fixed reimbursement for each patient on her panel. This reimbursement may be risk-adjusted so that the PCP will receive a higher reimbursement for less-healthy patients. Let  $r_m$  be the reimbursement that the PCP receives per year per patient in health state  $m$ . The PCP's *expected* reimbursement under a global capitation model,  $R_2(\mathbf{X}) = R_2 = \sum_m n_m r_m$ , is a constant.

In any global capitation setting, we equate the cost of prevention,  $v(\mathbf{X})$ , with the total capitation amount that that the PCP receives for its members. Therefore, in this case,  $v(\mathbf{X}) = R_2$ .

In a shared risk setting, the PCP's reimbursement is based on a *reference cost*,  $\tau$ , to the payer. If the total medical cost,  $\mathcal{Q}$ , exceeds the reference cost, the PCP is liable to the payer for  $\alpha(\mathcal{Q} - \tau)$ . If the total medical cost is less than the reference cost, the payer will pay the PCP a *bonus* in the amount  $\beta(\tau - \mathcal{Q})$ . Assume that  $0 \leq \alpha, \beta \leq 1$ . This liability/bonus payment is in addition to the reimbursement that the PCP receives for providing preventive care. Therefore, the PCP's 'gain' is given by  $G(\mathbf{X}) = \beta[\tau - \mathcal{Q}(\mathbf{X})]^+ + \alpha[\mathcal{Q}(\mathbf{X}) - \tau]^-$ . A negative gain is a loss to the PCP, i.e., an amount that the PCP must remit to the payer.

The PCP's *expected* reimbursement, in a fee-for-service *with shared risk* setting is given by  $R_3(\mathbf{X}) = R_1(\mathbf{X}) + G(\mathbf{X})$ . The PCP's *expected* reimbursement, in a global capitation *with shared risk* setting is given by  $R_4(\mathbf{X}) = R_2(\mathbf{X}) + G(\mathbf{X})$ .

We ignore fixed costs at this time. Suppose that the PCP incurs a cost of  $c_k$  per patient for intervention  $k$ . Her total costs are given by  $C(\mathbf{X}) = \sum_k c_k \sum_m n_m x_{km}$ . Given the PCP's intervention policy,  $\mathbf{X}$ , the PCP's profit,  $P(\mathbf{X})$ , is the difference between her reimbursements and her costs, i.e.,  $P(\mathbf{X}) = R(\mathbf{X}) - C(\mathbf{X})$ . Her decision problem is to select an intervention policy,  $\mathbf{X}$ , that will maximize profit, while satisfying capacity constraints. This problem is expressed as the following optimization problem.

$$\begin{aligned} \max_{\mathbf{X}} \quad & P(\mathbf{X}) \\ \text{s.t.} \quad & \sum_k \sum_m n_m d_k x_{km} \leq A \quad (1) \\ & 0 \leq x_{km} \leq 1 \quad \forall k, m \quad (2) \end{aligned}$$

$P(\mathbf{X})$  is non-linear in the cases where there is shared risk. It is possible to perform a linearization of the problem so that it can be solved using a mixed-integer linear programming solver such as CPLEX.

## REFERENCES

- Barhak, J., D. J. Isaman, W. Ye, and D. Lee. 2010. "Chronic disease modeling and simulation software". *Journal of Biomedical Informatics* 43 (5): 791–799.
- Bigus, J. P., C.-H. Chen-Ritzo, and R. Sorrentino. 2011, December. "A Framework for Evidence-Based Health Care Incentives Simulation". In *Proceedings of the 2011 Winter Simulation Conference*, edited

- by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu, 1103–1116. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Clarke, P. M. et al. 2004. “A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS 68)”. *Diabetologia* 43:1747–1759.
- Powell, W. B. 2007. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley Interscience.
- Sutton, R. S., and A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press.
- Watkins, C. 1989. *Learning from Delayed Rewards*. Ph. D. thesis, Cambridge University.

## **AUTHOR BIOGRAPHIES**

**JOSEPH P. BIGUS** is a Senior Technical Staff Member at the Thomas J. Watson Research Center, where he leads the ABLE research project. He is a member of the IBM Academy of Technology and is an IBM Master Inventor, with over 25 U.S. patents. He was an architect of the IBM Neural Network Utility and Intelligent Miner for Data products. Dr. Bigus received his M.S. and Ph.D. degrees in computer science from Lehigh University, his MBA from U. Mass., Amherst, and a B.S. in computer science from Villanova University. His current research interests include agent-based modeling and simulation and the application to understanding healthcare systems payments and incentives. His email address is [bigus@us.ibm.com](mailto:bigus@us.ibm.com).

**CHING-HUA CHEN-RITZO** is a Research Staff Member in the Business Analytics and Mathematical Sciences Department at the IBM T.J. Watson Research Center. She received a B.S. degree in physics from the University of North Carolina at Chapel Hill in 1997, an M.S. degree in Architectural Engineering in 1999, and a dual-title Ph.D. degree in Business Administration and Operations Research in 2006, both from Penn State University. She has been at IBM since 2005, where she has worked on applied mathematical optimization and simulation for applications in services and manufacturing operations. Her email address is [chenritzo@us.ibm.com](mailto:chenritzo@us.ibm.com).

**KEITH HERMIZ** is a Business Analytics Researcher at the IBM T.J. Watson Research Center. He has held a variety of positions at IBM including leadership of internal and client facing analytics consulting practices. He received his BS in Applied Mathematics from Brown University, MBA from the University of Rhode Island and Ph.D. from the University of Maryland at College Park. His email address is [khermiz@us.ibm.com](mailto:khermiz@us.ibm.com).

**GERALD TESAURO** is a Research Staff Member at IBM’s TJ Watson Research Center. He has worked on theoretical and applied machine learning in a wide variety of settings, including multi-agent learning, dimensionality reduction, computer virus recognition, computer chess (Deep Blue), intelligent e-commerce agents and autonomic computing. Most recently, he developed game-playing strategies for Watson, IBM’s Jeopardy! supercomputer. Dr. Tesauro received BS and PhD degrees in physics from University of Maryland and Princeton University, respectively. His email address is [gtesauro@us.ibm.com](mailto:gtesauro@us.ibm.com).

**ROBERT SORRENTINO** is a Research Staff Member in the Healthcare Transformation Department at the T.J. Watson Research Center. He received a S.B. degree in physics from MIT in 1973, and M.D. degree from Stony Brook University School of Medicine in 1978. He practiced Emergency Medicine for fifteen years at a high-acuity, inner-city hospital in Arizona. In addition to a long-term role as an independent healthcare management consultant for Mercer Health and Benefits, Dr. Sorrentino was Chief Medical Information Officer at CareMore Health Plan, and its MSO affiliate, CareMore Medical Enterprises, from 2002 through its acquisition by JPMorgan Partners in 2006. He subsequently served as Chief Medical Officer for ARTA Medicare Healthplan, and its affiliate MSO, Western Medical Management in Southern California. At IBM Dr. Sorrentino has worked on advanced healthcare analytics, and innovative payment models for healthcare. His email address is [sorrentino@us.ibm.com](mailto:sorrentino@us.ibm.com).