

THE REFERENCE MODEL ESTIMATES MEDICAL PRACTICE IMPROVEMENT IN DIABETIC POPULATIONS

Jacob Barhak

-

Austin, TX, USA

jacob.barhak@gmail.com

ABSTRACT

The Reference Model for disease progression is a league of disease models that validates simulation results to multiple observed clinical study results. Recent advances allows merging models, that were extracted from different populations in different decades, to be combined into one ensemble model that better validates against different clinical trials across years. Since each model and validation data set has a timestamp, it is possible to account for medical practice improvement during modeling. In the past it was observed that medical practice improvement caused models to become outdated and that adding temporal correction improved model fitness. However, it was unclear how much of the improvement is from prevention and how much from post event treatment. In this paper the rate of medical practice improvement is calculated through optimizing the model mixture considering diabetic populations. Results suggest similar improvement rates for prevention and post event treatment considering accumulated knowledge.

Keywords: Diabetes, Disease Progression, Optimization, Ensemble Models.

1 INTRODUCTION

Medical practice constantly improves as time passes and our lifespan increased in the last few centuries and decades. Especially notable is treatment in cardiovascular disease (CVD) in the last few decades. This improvement was reported by (Gregg et. al. 2012). Moreover, several well known risk equations were remodeled to fit more recent population data than it was originally modeled upon.

Since disease models try to describe phenomenon observed by clinical data, medical practice improvement should be incorporated within the model to account for model outdating.

So far modelers dealt with model outdating by periodically providing a new version of the model. Examples include:

1. The Framingham risk equation has been updated several times and several versions have been reported (Wilson et. al. 1998 and D'Agostino 2008).
2. The UKPDS diabetes outcome model has been updated recently (Hayes et. al. 2013 and Clarke et. al. 2004)
3. The QRisk equation is being updated regularly on a yearly basis (Hippisley-Cox et. al. 2008).

The Reference Model for disease progression showed that correcting older models for time passed since model data collection improves model fitness (Barhak 2014). However, the rate of improvement was not quantified and was assumed to be the same as reported in (Gregg et. al. 2012). The Reference Model was used in this paper to quantify the level of medical practice improvement based on 9 diabetic populations:

1. ASPEN (Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in Non-Insulin-Dependent Diabetes Mellitus)
2. ADVANCE (Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation)
3. ACCORD (Action to Control Cardiovascular Risk in Diabetes)
4. UKPDS (United Kingdom Prospective Diabetes Study)
5. KP (Kaiser Permanente)
6. NDR (Swedish National Diabetes Register)
7. Look AHEAD (Action for Health in Diabetes)
8. ADDITION (Anglo-Danish-Dutch Study of Intensive Treatment In PeOple With screenN Detected Diabetes in Primary Care)
9. CARDS (Collaborative Atorvastatin Diabetes Study)

The Reference Model was first constructed to validate multiple published models against multiple clinical trials. The idea was to accumulate clinical knowledge gathered by clinicians and models representing statistical assumptions in one location. At first, model combinations were competing against each other and High Performance Computing (HPC) was used to carry validations of many model combinations to many observed clinic trials outcomes.

Recent advances in modeling technology generated an assumption engine that constructs an ensemble model that allows models to cooperate or compete (Barhak, Garrett and Pruett 2016 and Barhak 2016). The assumption engine views models as assumptions and using optimization it deduces which assumptions better fit observed clinical trial data.

2 THE REFERENCE MODEL TECHNOLOGY

The Reference Model is currently composed of three main processes: Coronary Heart Disease (CHD), stroke, and competing mortality as shown in Figure 1. Each process is a separate state transition model where each process competes for reaching death during Monte Carlo micro-simulation. Transitions within each process occur with a probability that depends on the equation used. Each such equation can depend on multiple biomarkers defined for the individual. And those biomarkers can change during simulation. The most simple example is the age of the individual that increases by one each simulation step prior to calculating the transition probability.

The biomarkers are initialized at the beginning of simulation to match clinical trial populations distributions. The population generation process is itself a separate Monte Carlo simulation combined with evolutionary computation and discussed in (Barhak and Garrett 2014 and Barhak 2015). Population distributions are extracted from clinical trials reports. Those reports not only define starting population characteristics, they also define observed outcomes such as numbers of deaths observed in the trial period. The Reference Model is a validation model – it validates simulation results against those observed numbers and reports an error measure called fitness.

The fitness can be used as a scoring mechanism to quantify the ability of a certain equation to mimic the behavior of a certain clinical trial. Moreover, multiple validations can be performed to create a fitness score that quantifies the behavior of that equation for multiple populations. HPC is used to enhance this capability and quantifying the fitness of risk equation combinations. HPC allowed parallelizing fitness calculation of different populations and different risk equation.

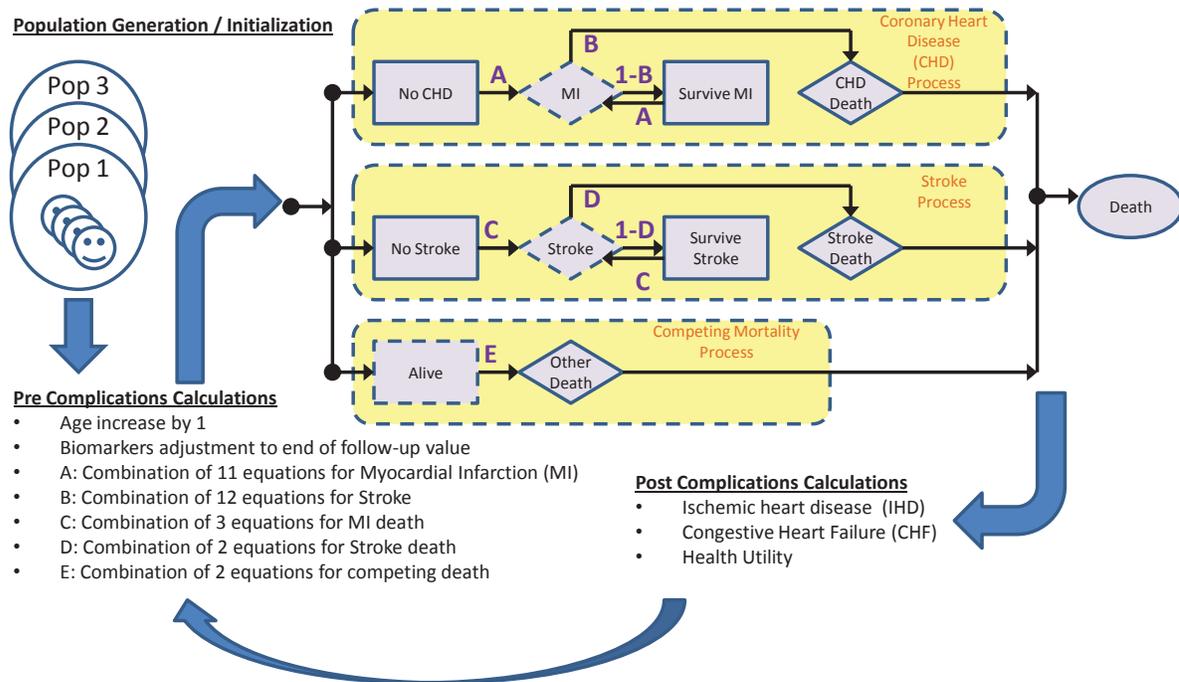


Figure 1: The Reference Model structure summary.

The result was a fitness matrix that acted like a scoring board in a sports league indicating scores of models and populations. The best model could be deduced by looking at the matrix. It is also possible to visually deduce which populations are easier and which are harder to explain using different models.

The fitness matrix represented discrete model space where model combinations were competitive, relying on equation swaps for the same transition probability. So if a transition probability could be represented by two equations A_1 and A_2 , then the system had to choose only one of them. Recently this was changed in (Barhak 2016) to allow merging both equations in a linear combination such as $\omega_1 A_1 + \omega_2 A_2$ using coefficients ω_1 and ω_2 . Therefore the discrete simulation matrix was replaced by a continuous model space represented by these coefficients. This redefinition allowed using optimization techniques used in machine learning to distill the best model using the fitness score. This technique was named: “assumption engine”.

3 THE ASSUMPTION ENGINE EXTRACTS MEDICAL PRACTICE IMPROVEMENT

The assumption engine got its name since the modeler can “throw” assumptions at it and the assumption engine can make sense out of them. Since models are assumptions of how reality behaves, the assumption engine is very useful for decision makers to figure out which assumptions make sense in the context of the data and query. The assumption engine will highlight equations that fit the data by assigning them a large coefficient and will reject equations that generate poor fitness by diminishing their coefficient.

At the heart of the assumption engine is the gradient descent algorithm, yet there are several mechanisms that keep the optimization process on track and avoid unreasonable results:

1. Coefficients can be bounded during optimization to avoid situations such as a negative coefficient for an equation, indicating that the equation subtracts rather than contributes information. This is unreasonable since modelers work hard and report reasonable equations and do not attempt to

mislead. If a risk equation reports a number that does not validate well, we want to ignore it rather than use its value to subtract from another equation.

2. Coefficient groups can be scaled to sum a specific number. This is useful during optimization to keep equation combinations within the convex hull of modeling space. This will ensure that the combined probability is never higher than the highest probability or lower than the lowest probability that is reported by any risk equation.
3. Several initial guesses are provided to the assumption engine and optimized in parallel towards best fitness. Moreover, the gradient components are numerically computed in parallel using HPC. Parallelization reduces computation time, yet more importantly it allows for testing multiple assumptions and strives towards finding best global fitness rather than local fitness.

The first two elements also ensure that event probability is never above 1 or below zero. without those constraints, risk equation combination could have resulted in an invalid probability, and those constraints keep the simulation mathematically valid.

Yet despite the third parallelization element that seeks global minimum during optimization, there may be multiple other solutions not explored, after all the modeling space is infinite now. Moreover, since the Monte Carlo simulation generates an error, the optimization process has limited accuracy. Nevertheless, sufficient repetition powered by HPC improves the combined model compared to each of the original risk equations uncombined. For example we can create a new risk equation that averages two others, one overestimating and another underestimating.

An important observation in the context of this paper is that coefficients can represent assumptions that supplement equations and not only equation significance coefficients. It allows us to construct new models from computational components that include assumptions. In this paper it allows us to correct models to account for medical progress.

4 ACCOUNTING FOR MEDICAL PRACTICE IMPROVEMENT

Any model should be viewed as an assumption the modeler made on how reality behaves according to observations. With this point of view in mind, a risk equation is actually a compressed form of the data it is derived from. Therefore it implicitly embeds within it characteristics such as location and time that the data represents. In this paper we consider the temporal aspect of each equation as a modifying factor.

For example, a risk equation describing mortality in the middle ages it clearly not appropriate to predict lifespan of modern society. This negative example indicates that some correction is appropriate. The solution presented in this paper is to account for time passed since equation formation time to equation application time.

Each risk equation used is therefore assigned a timestamp that represents the midpoint between study start and study end. Since baseline cohort collection time spread is recorded differently and since currently clinical trial reports are textual and not come in machine readable form, the calculation may vary due to deduction effort and assumptions. Yet regardless of exact calculation, the idea is that the timestamp associated with a risk equation is roughly the center point of the trial. For example UKPDS reports recruitment dates of 1977-1991 with 10.7 follow-up according to (Stevens and Kothari 2001). Therefore, the model timestamp was calculated as $(1977+1991+10.7)/2=1989.35$. Model timestamps for the 9 populations used in this paper ranged from 1978.5 to 2007.05. From here on this timestamp for individual i will be denoted as y_i .

The Reference Model applies each equation to an individual that also has a timestamp. This timestamp is calculated during generation of each individual as a uniform distribution between recruitment start to recruitment end. For example, UKPDS population was distributed uniformly between 1977 to 1991 as reported in (UKPDS 1988). Then each simulation time step, the individual timestamp is advanced by 1,

we will denote this timestamp for risk equation j as y_j^{Model} . However, since the model is an ensemble model built from risk equations with coefficients ω_j the timestamp of the ensemble model is:

$$\hat{y}^{Model} = \frac{\sum_j \omega_j y_j^{Model}}{\sum_j \omega_j}$$

Having two timestamps for the model and the individual allows accounting for differences in time passed between the model data was observed and its application. An exponential correction term was selected to affect the transition probabilities for Myocardial Infarction (MI), MI death, stroke, and stroke death. The correction term is:

$$p^{corrected} = \mu^{\varphi\tau} p$$

Where p represents the yearly transition probability that is being corrected, and τ represents the difference between model and individual timestamps in years

$$\tau = y_i - \hat{y}^{Model}$$

The coefficient μ represents the cardiovascular disease improvement coefficient for one year difference in timestamp between the model and the individual. Cardiovascular disease improvement includes MI and stroke and both prevention and post event treatment improvement.

The coefficient φ changes between prevention affected probabilities such as MI and stroke and post event treatment probabilities such as MI mortality and stroke mortality.

$$\varphi_{prevent} = \varphi = 1 - \varphi_{post_event}$$

The prevention coefficient φ is complimentary to the post event coefficient $1 - \varphi$. This means that when multiplying the prevention and post event probabilities, the result would be dependent only on the coefficient μ and the time assuming that the timestamp difference is the same for the equations:

$$\begin{aligned} p_{Death}^{corrected} &= p_{Healthy\ to\ Event}^{corrected} * p_{Event\ to\ Death}^{corrected} \\ &= \mu^{\varphi\tau} p_{Healthy\ to\ Event} * \mu^{(1-\varphi)\tau} p_{Event\ to\ Death} \\ &= \mu^{\tau} p_{Healthy\ to\ Event} * p_{Event\ to\ Death} \end{aligned}$$

And do note that this is a simplification and an arbitrary construct, another formulation of this modeling assumption may generate another mathematical form that will behave differently. For example, an alternative approach would have been to just assign a temporal correction coefficient for each transition probability rather than the construct shown above that couple transition probabilities. The chosen formulation is a simplification largely affected by historical reasons to compare to the number reported by (Gregg et. al. 2012).

For historical development reasons the first coefficient is implemented as a 6 year coefficient similar to (Gregg et. al. 2012) that reported improvement for 6 years. So the actual simulation reports the 6 year coefficient rather than a yearly improvement coefficient, yet in this paper numbers are translated to a yearly coefficient for clarity. The relationship is:

$$\mu = \mu_{Six\ Years}^{(1/6)}$$

This simplified correction term was applied to all transition probabilities other than competing mortality. Recall that a transition probability is composed from several equations with multiple coefficients. These coefficients will be optimized together with the temporal correction coefficients. This is a major change from previous work (Barhak 2014) where the coefficient μ was set the (Gregg et. al. 2012) value of 0.6

and φ was set to a value that historically improved fitness after trial and error. With the introduction of the assumption engine it was possible to efficiently calculate the temporal coefficients when considering a model mixture. In this work, both μ and φ are byproducts of optimization from data collected. In other words the assumption engine was asked to deduce the best model mixture taking into account correction for model outdateding in its results.

5 RESULTS

There were two simulations conducted for this paper:

1. Control: optimizes the best model without temporal correction
2. Temporal Correction: optimizes the best model together with temporal correction.

In both simulations conditions were similar with one major difference: in the control simulation there was no temporal correction applied to the simulation and in the temporal correction simulation, temporal coefficients were applied to CVD transition probabilities as explain in the previous section.

In both cases simulations included validation against 47 cohorts from 9 populations: ASPEN, ADVANCE, ACCORD, UKPDS, KP, NDR, Look AHEAD, ADDITION, CARDS. Initial conditions were used with and without biomarker correction. Biomarker correction changed the biomarker value to end of trial values when reported, indicating the treatment effect for A1C, BMI, smoking, blood pressure, and cholesterol. The initial guess for the gradient descent algorithm in the assumption engine assumed that all risk equations have a similar contribution for each category, i.e. coefficients value for each category is $\omega_j = 1/N$ where N is number of equations in that category. For example coefficients associated with stroke death were initialized to $\omega_j = 1/2$ since there are only 2 stroke death equations used. There were 11 MI equations, 12 Stroke equations, 3 MI death equations, 2 Stroke death equations, 2 competing mortality equations.

Results are displayed in Table 1 and in Figure 2. The first column represents the element calculated in optimization where risk equations correspond to the transitions in Figure 1 marked by the letters A,B,C,D,E. MI Eq. contribute to transition A. MI Death Eq. contribute to transition B. Stroke equation contribute to transition C, Stroke death equations contribute to transition D, and Death equations correspond to transitions marked with E in Figure 1. Table rows show the best result obtained in optimization after 10 iterations. For equations the numbers corresponds to the coefficient of a certain equation used in the model mixture – these are the ω_j values after optimization. This is how much the equation contributes to the mixture of equations – for simplicity this can be described as the weight of an equation.

Note that only unperturbed results are displayed, i.e. simulations conducted to calculate perturbed gradient calculate gradient components for gradient descent were not included in the paper, although some report better fitness. A quick look at the coefficients of the risk equations reveal that some risk equations lose their importance with time passing, yet once corrected for time passing they regain their importance, see MI equation 1 for example that is not significant in the control simulation yet becomes valuable in the temporal corrected simulations. The results also show that most fluctuations in coefficient in the model mixture are for MI and stroke equations while death equations used seem to be less prone to coefficient change in the model mixture. At the bottom of the tables we include convergence information of best fitness reached and the gradient descent iteration that this best result was reached.

Recall that the purpose of optimization was to improve the fitness score. This score was improved significantly from 38.39 to 16.77 when using biomarker correction or from 44.46 to 16.92 when biomarkers are steady throughout simulation. Figure 2 visually demonstrates this dramatic improvement in the topmost bar set. This clearly indicated that models should include temporal correction. Moreover, near the bottom of Table 1 and near the top of Figure 2 the results also estimate the yearly improvement coefficient and the prevention and post event improvements which are the focus of this paper.

Table 1: The assumption engine optimization results.

Parameter	Temporal + Bio- Marker	Temporal	Control + Bio-Marker	Control
MI Eq. 1	0.32	0.33	0.05	0.07
MI Eq. 2	0.00	0.05	0.10	0.02
MI Eq. 3	0.02	0.07	0.07	0.02
MI Eq. 4	0.40	0.31	0.00	0.09
MI Eq. 5	0.00	0.09	0.15	0.21
MI Eq. 6	0.09	0.10	0.32	0.16
MI Eq. 7	0.00	0.00	0.09	0.07
MI Eq. 8	0.17	0.00	0.06	0.28
MI Eq. 9	0.00	0.03	0.02	0.02
MI Eq. 10	0.00	0.02	0.14	0.06
MI Eq. 11	0.00	0.00	0.00	0.00
Stroke Eq. 1	0.40	0.26	0.13	0.12
Stroke Eq. 2	0.26	0.12	0.25	0.07
Stroke Eq. 3	0.00	0.10	0.17	0.07
Stroke Eq. 4	0.01	0.09	0.22	0.12
Stroke Eq. 5	0.00	0.00	0.03	0.12
Stroke Eq. 6	0.00	0.04	0.00	0.04
Stroke Eq. 7	0.29	0.01	0.06	0.04
Stroke Eq. 8	0.04	0.02	0.06	0.06
Stroke Eq. 9	0.00	0.06	0.01	0.11
Stroke Eq. 10	0.00	0.04	0.04	0.14
Stroke Eq. 11	0.00	0.08	0.04	0.02
Stroke Eq. 12	0.00	0.17	0.00	0.10
Death MI Eq. 1	0.26	0.41	0.26	0.33
Death MI Eq. 2	0.35	0.25	0.37	0.39
Death MI Eq. 3	0.38	0.34	0.37	0.28
Death Stroke Eq. 1	0.45	0.51	0.43	0.49
Death Stroke Eq. 2	0.55	0.49	0.57	0.51
Death Eq. 1	0.53	0.47	0.57	0.57
Death Eq. 2	0.47	0.53	0.43	0.43
Prevention Coefficient	0.43	0.53	N/A	N/A
CVD 6 Yr Coefficient $\mu_{six\ years}$	0.41	0.44	N/A	N/A
Yearly CVD Coefficient μ	0.86	0.87	N/A	N/A
Optimization Iteration Results was Reached	9	10	8	9
Fitness	16.77	16.92	38.39	44.46

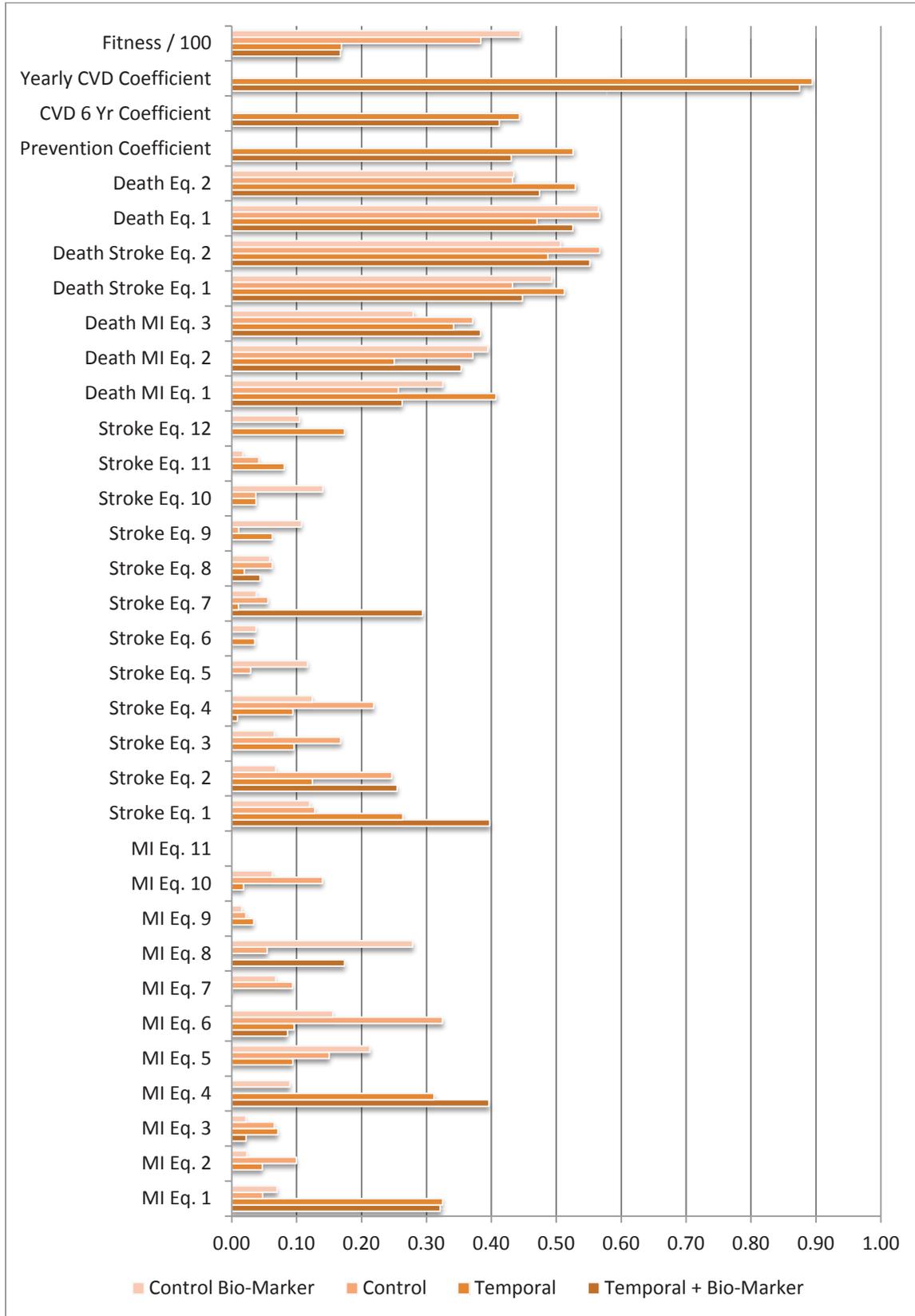


Figure 2: The assumption engine optimization results.

6 DISCUSSION

The results presented here clearly indicate that there was a major improvement in treating CVD in diabetic populations in the last few decades. The yearly CVD coefficient μ was calculated as 0.86/0.87 in the simulations that include populations with timestamps ranging from 1978.5 to 2007.05. The results also suggest that the improvement was both due to prevention and post event treatment with roughly similar quantities indicated by the prevention coefficient φ of 0.43/0.53. Notice that different initial conditions reach slightly different coefficients for improvement depending on model formulation. The combined model that includes biomarker correction indicates slight preference to medical improvement due to post event treatment while the models without biomarker correction indicate that improvement was slightly better for prevention. Yet differences should not be considered significant without further validation since there is much more uncertainty as discussed below. However, the numbers do represent a ballpark range that should not be ignored.

Moreover, these results correspond to those previously reported in (Gregg et al. 2012) that reported $\mu_{Six\ Years}$ of 0.6 (0.46 - 0.77). There is a difference between the level of improvement predicated by the simulation of 0.41/0.44 for both scenarios – this is more optimistic than what was calculated in (Gregg et al. 2012). Even though exact estimates of improvement impact are not the same which is reasonable considering different time periods addressed, the idea that medical practice improved has been reassured once more with different methods based on modeling.

There can be many reasons why numbers can differ:

- 1) The data used for validation includes noise. For example, if one repeats the same clinical trial there will be some statistical fluctuations in reported results. This is diminished since many numbers from multiple trials are being compared to, yet even the best data contains noise.
- 2) Monte Carlo Simulation noise – the simulation conducted here were executed for groups of 1000 individuals sampled from each of the 47 cohorts representing the 9 populations used. So noise is present and may affect results. Do note that there may be multiple mathematical representations that will produce best fitness as described in (Barhak, Garrett and Pruett 2016) – so Monte Carlo noise may have additional affect. However, the simulations took roughly a year of computation, reduced to about 3 weeks on a 16 core cluster. It is possible to reduce this error further by investing more computing power into the computation, yet considering the resources available and fluctuations in results between gradient components, results seem reasonable. After all, fitness was considerably improved during optimization from control and this improvement was consistent – the worst temporally corrected model in the last iteration improved by more than half in fitness compared to the control best model.
- 3) The initial guess introduced to the system was somewhat arbitrary and the assumption engine will converge to a local minimum of fitness. There may be multiple minima that may produce different results. However, since there were 2 competing scenarios that generated similar temporal improvement coefficients, and reached considerable improvement compared to the control, these results are probably representative of the trend indicating the need for temporal correction in models.
- 4) Gradient descent iterations were capped at 10 and optimization step sizes were preset to constant values. It is possible that a local minimum has not been reached and additional time would have resulted in a better model. Yet due to computing power constraints and the significant improvement in fitness achieved already by temporal correction it seemed reasonable to keep those results as these are unlikely to reach much better fitness.
- 5) The assumption engine accumulates knowledge in the form of assumptions represented by models and observed data depicted by populations. If more data will be accumulated, the result may shift to fit the newer data. And do recall that a model is an assumption the modeler made, therefore if the model includes elements of treatment that behave better across time, then it will change the temporal improvement coefficients. The accumulation of data is an iterative process.

As those words are written a much larger simulation is underway that roughly doubles the number of populations validated against. So the results here should be viewed as best possible to date within limitations. This is a mere milestone that future work can be compared to.

- 6) Human error should be considered since the system still relies on elements handled by humans. For example, after the above results were obtained it was discovered that two numbers in CVD outcomes in the ACCORD trial were switched between two cohorts. Since the optimization is towards 224 numbers, those accidentally switched numbers should not change results significantly beyond other errors. And when many numbers are reported, human entry errors should be expected. The way to reduce such uncertainty is accumulating information. With sufficient amount of data accumulated, the noise due to human errors is reduced. As time passes the amount of these errors reduces since they are discovered and eliminated by repeated testing conducted when updating model version.
- 7) Dates associated with population data and model date were sometimes estimated from pieces of information rather than reported. For example, clinical trials report start and end dates of recruitment, yet do not report a temporal center point considering the entire timeline including follow-up time, so these have to be estimated. Moreover, the ensemble model date is an artificial mathematical construct and other forms of it are possible. Yet more importantly, models published do not report an exact date for which the model is applicable so estimates deduced from text may have inaccuracies. Since multiple models and data sets were used, this concern is less likely to impact results, yet it would help if such timestamps of center of mass time would be published in the future for models and reported in clinical trials as a standard.

The last argument is very important from a modeling perspective and not only from a clinical trial reporting perspective. Modelers not only have to publish meta data such as location and time along with their model, they also need to consider future application of the model. If a model aims to predict a distant future, it has to take into account future changes in modeling conditions. In medical modeling, where treatments and guidelines regularly change, modelers should incorporate assumptions of future medical changes as part of the model. Otherwise models become outdated and the modeling effort needs to be repeated every few years.

The effect of medical practice improving in long term predictions using the best model mixture deduced in this paper is discussed in details in (Barhak 2017) and although strongly connected to this paper and relies on results presented here, it is outside the scope of this paper. The reader is encouraged to read this continuation paper to comprehend the impact of the results presented here.

7 REPRODUCIBILITY INFORMATION

The results for this paper were calculated on a 16 core cluster with 5 nodes running Ubuntu 12.04 Linux using Sun Grid Engine and Python 2.7.8 deployed by Anaconda 2.0.1 (64-bit). The Reference Model results were generated using MIST version (0,94,2,0) with Inspyred version 1.0. Results are archived in: MIST_RefModel_2016_10_26_OPTIMIZE.zip for temporal correction using model version 44 and in MIST_RefModel_2016_12_17_OPTIMIZE.zip for control using model version 45. Numbers appearing in paper were rounded for display purposes. The results displayed are based on models optimized towards an outcome that contained a switch between CVD outcomes in two ACCORD trial cohorts. This error was discovered after results were computed and is reported here for future reference in case a third party tries to reproduce this work.

ACKNOWLEDGMENTS

Many thanks to Aaron Garrett and W. Andrew Pruett who helped develop the mathematical technique that helped power this work. Thanks to Chad Ridgley for proof reading the paper. This paper used the Micro Simulation Tool (MIST) that is based on IEST. The IEST GPL disease modeling framework was

initially supported by the Biostatistics and Economic Modeling Core of the MDRTC (P60DK020572) and by the Methods and Measurement Core of the MCDTR (P30DK092926), both funded by the National Institute of Diabetes and Digestive and Kidney Diseases. The modeling framework was initially defined as GPL and was funded by Chronic Disease Modeling for Clinical Research Innovations grant (R21DK075077) from the same institute. The Reference Model for Disease Progression and the MIST framework were developed without financial support.

REFERENCES

- Barhak J. 2014, “The Reference Model for Disease Progression uses MIST to find data fitness”. *PyData Silicon Valley 2014*: Available via: <https://www.youtube.com/watch?v=vyvxiljc5vA> Presentation: http://sites.google.com/site/jacobbarhak/home/PyData_SV_2014_Upload_2014_05_02.pptx Accessed Dec. 23, 2016.
- Barhak J. 2015. “The Reference Model uses Object Oriented Population Generation”. *SummerSim 2015*. Chicago IL, USA. Available via: <http://dl.acm.org/citation.cfm?id=2874946> Presentation retrieved from: http://sites.google.com/site/jacobbarhak/home/SummerSim2015_Upload_2015_07_26.pptx
- Barhak J. 2016. “The Reference Model for Disease Progression Combines Disease Models”. *IITSEC 2016*. Available via: <http://www.iitsecdocs.com/volumes/2016> Presentation: http://sites.google.com/site/jacobbarhak/home/IITSEC2016_Upload_2016_11_05.pptx Accessed Dec. 23, 2016.
- Barhak J. 2017. “The Reference Model Predicts Life Expectancy for Diabetics”. *MODSIM World 2017*. Apr 25-27, 2017, Virginia Beach, VA.
- Barhak J., Garrett A. 2014. “Population Generation from Statistics Using Genetic Algorithms with MIST + INSPYRED”. *MODSIM World 2014*, Available via: http://sites.google.com/site/jacobbarhak/home/MODSIM2014_MIST_INSPYRED_Paper_Submit_2014_03_10.pdf Presentation: http://sites.google.com/site/jacobbarhak/home/MODSIM_World_2014_Submit_2014_04_11.pptx Accessed Dec. 23, 2016.
- Barhak J., Garrett A., and Pruett W. A. 2016. “Optimizing Model Combinations”, *MODSIM world 2016*. Available via: http://www.modsimworld.org/papers/2016/Optimizing_Model_Combinations.pdf Presentation: http://sites.google.com/site/jacobbarhak/home/MODSIM2016_Submit_2016_04_25.pptx Accessed Dec. 23, 2016.
- Clarke P.M., Gray A.M., Briggs A., Farmer A.J., Fenn P., Stevens R.J., Matthews D. R., Stratton I. M., Holman R. R., & UK Prospective Diabetes Study (UKDPS) Group. 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 no. 68). *Diabetologia*, Vol. 47(10), pp. 1747-59. <http://dx.doi.org/10.1007/s00125-004-1527-z>
- D'Agostino R.B., Vasan R.S., Pencina M.J., Wolf P.A., Cobain M., Massaro J.M., Kannel W.B. 2008. "General cardiovascular risk profile for use in primary care: the Framingham Heart Study". *Circulation* Vol. 117 (6) pp. 743–753. <https://dx.doi.org/10.1161%2Fcirculationaha.107.699579>
- Gregg E. W., Cheng Y. J., Saydah S., Cowie C., Garfield S., Geiss L., and Barker L. 2012. “Trends in Death Rates Among U.S. Adults With and Without Diabetes Between 1997 and 2006 - Findings from the National Health Interview Survey”. *Diabetes Care* Vol. 35(6) pp. 1252–1257. <https://doi.org/10.2337/dc11-1162>
- Hayes A.J., Leal J., Gray A.M., Holman R.R., & Clarke P.M. 2013. “UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82”. *Diabetologia*, Vol. 56(9), pp. 1925-33. <http://dx.doi.org/10.1007/s00125-013-2940-y>

- Hippisley-Cox J., Coupland C., Vinogradova Y., Robson J., Minhas R., Sheikh A., Brindle P. 2008. "Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2". *BMJ* Vol. 336, pp. 26, doi: <http://dx.doi.org/10.1136/bmj.39609.449676.25>
- Stevens R., Kothari V., Adler A., Stratton I. 2001. "The UKPDS risk engine: a model for the risk of coronary heart disease in type II diabetes UKPDS 56". *Clinical Science* Vol. 101 pp. 671-679.
- UKPDS: UK Prospective Diabetes Study UKPDS Group. 1988. "Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes UKPDS 33." *Lancet* Vol. 352 pp. 837–853, [http://dx.doi.org/S0140-6736\(98\)07019-6](http://dx.doi.org/S0140-6736(98)07019-6)
- Wilson P. W. F., D'Agostino R. B., Levy D., Belanger A. M., Silbershatz H. and Kannel W. B. 1998, "Prediction of Coronary Heart Disease Using Risk Factor Categories". *Circulation* Vol. 97 pp. 1837-1847. <https://doi.org/10.1161/01.CIR.97.18.1837>

AUTHOR BIOGRAPHIES

JACOB BARHAK specializes in chronic disease modeling with emphasis on using computational technological solutions. Dr. Barhak has diverse international background in engineering and computing science. He received his Ph.D. in 2003 from the Technion Israel Institute of Technology. From 2003 to 2006 he worked at an NSF Engineering Research Center at the University of Michigan focusing on inspection systems and personalization. From 2006 to 2012 he worked on developing disease modeling tools for the Michigan model for diabetes at the University of Michigan. The Reference Model for disease progression was independently self-developed by Dr. Barhak in 2012. He is the developer of the MICO Simulation Tool (MIST). His web page is: <http://sites.google.com/site/jacobbarhak/>