Identification of mechanistic differences in a virtual population contributing to variation in response to simulated statin therapy

Karim Wahba[1], Tom Paterson[1], and Lyn M Powell[1]

[1]*Entelos, CA, USA*

Correspondence: wahba@entelos.com; +1 (650) 572 5436; 2 Waters Park Drive, Suite 200, San Mateo, CA 94403, USA

Introduction

The use of mechanistic *in silico* modelling of integrated biological processes at a system-wide level is a growing discipline that can revolutionize the field of medicine. The realization of this potential will transform the process of therapeutic discovery and development to better align with that of modern, complex engineering products: iterative cycles of hypothesis and simulation-driven testing processes. Moreover, integrating disparate biological data, in a consistent manner, to constrain model behaviours and to create virtual patient phenotypes will be essential for delivering on the promises of personalized medicine. For any predictive modelling approach, a major consideration is how to incorporate variability seen in the data resulting from heterogeneity in the underlying physiology, genetic diversity, and lifestyle/environmental factors.

Here we showcase the cardiovascular (CV) PhysioLab® and associated tools that were developed and used for physiological modelling at Entelos®. The CV PhysioLab is a mechanistic, large-scale dynamical model of cholesterol metabolism, atherosclerosis, and CHD risk. We focus here on the components associated with cholesterol homeostasis (Fig.1.), explaining how a top-down mechanistic modelling approach incorporates *in vitro* and *in vivo* data, and constraints on ranges for physiological parameters. The concept of a virtual patient (VP) is introduced as a parameterization of the model that yields a specific phenotype – for example, dyslipidemics who are hyper- or hypo- responsive to HMGCoA inhibitiors (statins). A virtual population (VPop) is created to match reported lipoprotein values of patients in the Treat to New Targets (TNT) trial [1], and their response to atorvastatin 10 mg qd. The creation of a diverse VPop enables the interrogation of the different mechanisms that distinguish hyper- and hypo- statin responders. The potential for identifying predictive biomarkers using this approach is highlighted here, as well as in peer-reviewed journal articles using other Entelos models such as the Metabolism PhysioLab, and the Rheumatoid Arthritis PhysioLab [2, 3].

The cholesterol metabolism submodel of the CV PhysioLab platform

The progression of atherosclerosis is hypothesized to be primarily driven by the balance between cholesterol retention and efflux from the vessel. This balance is dependent on the retention of circulating apolipo-protein (apo) B100/LDL particles within the plaque and cholesterol efflux from the plaque and other tissues to apo- AI /HDL particles. The cholesterol metabolism submodel computes the dynamic equilibrium of apo B100 and apo A-I species in the circul-ation, in terms of both particle numbers and cholesterol content. The model outputs plasma lipoprotein values relevant to clinical measures such as total cholesterol (TC) and

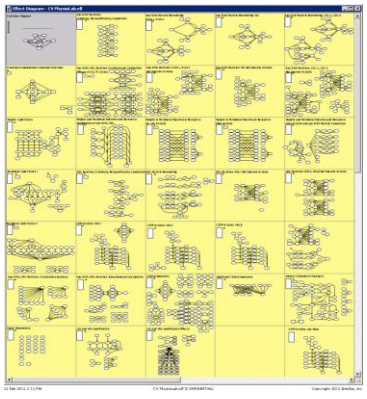

Figure 1. Overview of cholesterol submodel. The nodes and arrows correspond to state variables and relationships respectively

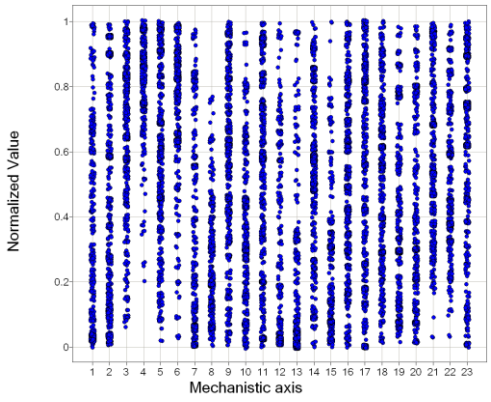

Figure 2. Coverage of mechanistic axes by 1600 VPs. In total, 23 pathways associated with lipoprotein metabolism are explored. The range for each pathway is normalized between 0 and 1.

triglycerides (TG), and various sub-particle classes such as VLDL, IDL, LDL and HDL. Activity of key enzymes (CETP, LPL, HL, LCAT, PCSK9) and key receptors and transporters (LDLr, SR-B1, ABCA1) which contribute to processes such as synthesis, catabolism and remodelling of particles are included. The entire submodel is based on the conservation of mass balance, and incorporates known feedback mechanisms. Inter-individual variability in VPs arises due to variations in the mechanisms (Fig.2.) resulting in individual lipoprotein profiles and differing responses to therapeutic interventions (Fig.3.). In total, 23 mechanistic pathways or "axes" are explored (e.g. LPL activity on VLDL particles, apo B100 synthesis rates, PCSK9 half-life, etc). The effects of cholesterol interventions (e.g. statins, CETP inhibitors, cholesterol absorption blockers etc.) are modelled as perturbations to the appropriate pathways which result in changes to the lipoprotein values consistent with clinical data.

Exploring model parameters and creating virtual patients and populations

The generation of VPs is done through use of a genetic algorithm to explore the parameter space to generate feasible patients. For a given axis, the range within which the search is performed is derived from the literature, and can utilize *in vitro* and *in vivo* human and/or animal data. For example, $VLDL_1$ apoB production in human can range from 3 – 16 mg/kg/day [4]. The drivers of VLDL synthesis rates in the model are, therefore, sampled to cover this ~5 fold range. A VP is considered feasible if their computed VLDL cholesterol falls within the range reported in the literature. Using this methodology, VPs representing different model parameterizations are created to reproduce feasible clinical phenotypes and responses (Fig.3.). This is a process that often requires multiple iterations to cover the full desired range of measurements. VPs of a specific phenotype can be used to seed more explorations, and their progeny enrich the phenotypic and mechanistic space, leading to a diverse cohort, in this example, consisting of >1600 VPs.

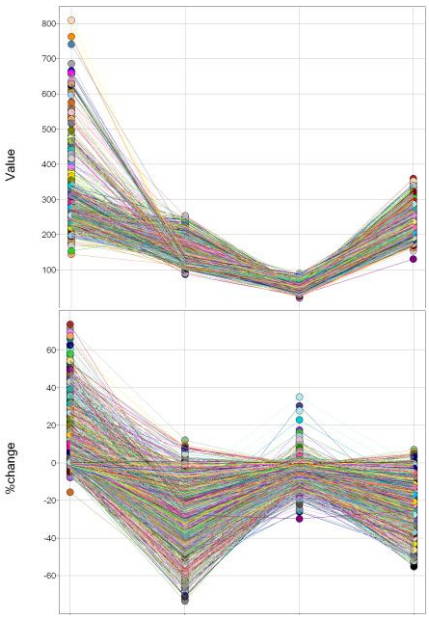

Figure 3. Each connected line is a profile of a VP (N = 1600) for 4 measures : PCSK9 (ng/ml) , LDL-C, HDL-C and TC (mg/dl) at baseline (top chart) and % changes in response to simulated atorvastatin 10 mg qd (bottom chart)

To adjust for any sampling bias of the feasible VPs returned by the genetic algorithm, and to ensure that the statistics of the generated VPs match clinical data, a prevalence weighting algorithm is used to assign a weight to each VP. This weight corresponds to the fractional contribution of a VPs phenotype to a specified population. Individual VP weights can be optimized to simultaneously match reported sample statistics such as means, standard deviations, distributions, and any known correlations, across multiple measures. In this test case, a cohort of >1600 VPs was weighted to match reported TNT means and standard deviations of LDL-C, HDL-C, TC and TG at baseline (washout), as well as LDL-C distributions in response to simulated atorvastatin 10 mg qd (Fig.4.). In particular, VPs with LDL-C <130 mg/dL or > 250 mg/dL after washout are excluded. The VPs corresponding to the top and bottom deciles of LDL-C response to statin (fractional change) are selected for further analysis.

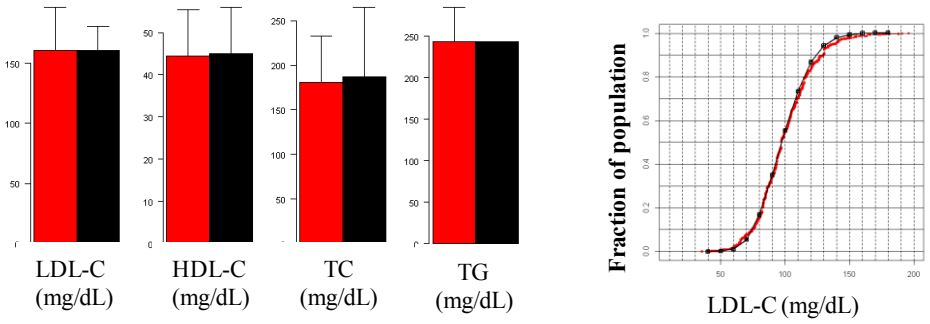

Figure 4. Comparison of simulated Vpop lipoprotein measures (red) to those reported for TNT trial (black). The left plot shows means and standard deviations at baseline (washout) and right plot shows cumulative distribution for LDL-C on atorvastatin 10mg.

Exploring mechanistic pathways for statin hypo- and hyper-responders in VPop

The mechanistic pathways that were explored in generating the VPs can be ranked to give a qualitative indication of their relative importance in distinguishing between hypo- and hyper- statin responders in the VPop, defined as the top and bottom deciles of fractional LDL-C response. One method of ranking these pathways is by looking for separations in the distributions of these two phenotype groups, or subpopulations within the VPop, across each mechanistic axis used in the model. A particular pathway may be implicated if there is a significant difference in probabilities. The plot in Fig. 5 shows the results for our TNT VPop. Pathways in the model that are highlighted by the arrows show qualitative differences between the two subgroups. In general, poor statin response may be associated with higher PCSK9 levels, lower LDL receptor synthesis rates, differential CETP activity etc.

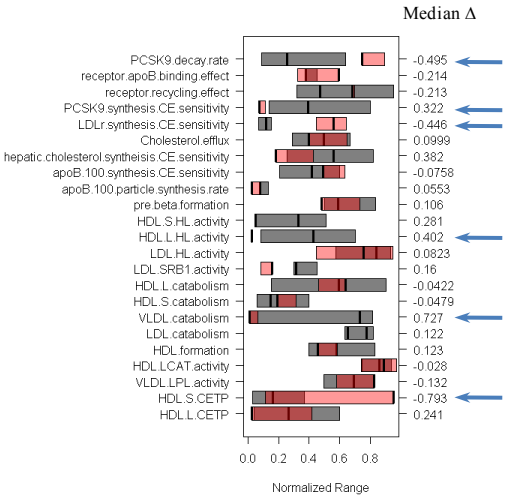

Figure 5. Distributions of top and bottom 10% of Vpop on statin for fractional LDL-C change across each of the 23 mechanistic pathways. Statin hyper-responders, in red, and hypo-responders, in grey, are plotted (median and IQR) on a normalized range for each mechanism. Dark red denotes an overlap of the two subpopulations. Pathways where the medians are more separated and with little overlap may be associated with distinguishing the two groups.

Conclusion

Mechanistic modeling of biological systems offers the opportunity to test hypotheses, and thereby manage risk, during multiple stages of the drug discovery process. Models that capture behaviour at the system-wide level, and which employ virtual patients and populations to explore uncertainty and diversity in underlying mechanisms are at a distinct advantage.

References

1. Waters D, Guyton J, Herrington D, McGowan M, Wenger N, Shear C, TNT Steering Committee Members and Investigators. Treating to New Targets (TNT) Study: does lowering low-density lipoprotein cholesterol levels below currently recommended guidelines yield incremental clinical benefit? Am. J. Cardiol. 2004; 93:154-158.

2. Klinke, D. Integrating epidemiological data into a mechanistic model of type 2 diabetes: validating the prevalence of virtual patients. Ann. Biomed. Eng. 2008;36:321-334.

3. Meeuwisse C, van der Linden M, Rullman T, Allaart C, Nelissen R, Huizinga T, *et al*. Identification of CXCL13 as a marker for rheumatoid arthritis outcome using an in silico model of the rheumatic joint. Arthritis Rheum. 2011;63:1265-1273.

4. Adiels M, Taskinen M, Packard C, Caslake M, Soro-Paavonen A, Westerbacka J, *et al*. Overproduction of large VLDL particles is driven by increased liver fat content in man. Diabetologia. 2006;49:755-765