Enabling Model Complexity Through an Improved Workflow

Christine Allen, James Collins, Zane Rankin, Kate Wilson, Derrick Tsoi, Kelly Compton, and Abraham D. Flaxman

Institute for Health Metrics and Evaluation, University of Washington

Abstract

Using data from the Global Burden of Disease study, the Institute for Health Metrics and Evaluation's Simulation Science team created a framework and modeling process to enable development of modular healthcare intervention packages, variable groupings of customizable intervention microsimulations. This allows for more timely and efficient model development, and provides the potential for greater model complexity. The team accomplished this by expanding upon their existing simulation framework, then identifying three specific process improvement goals: a method for specifying intervention packages; a modeling process that increases productivity; and a documentation strategy that facilitates transparency. They met those goals by continuing development of their existing tool, developing a clearly defined and iterative model workflow, and integrating within that workflow a clear set of templated phase outputs.

1   Introduction

First published in 1993, the Global Burden of Disease (GBD) study is the most comprehensive global epidemiological effort in history.[1] Based on extensive reviews of international literature, the GBD produces metrics for more than 300 diseases in 195 countries. The study estimates metrics of disease burden such as prevalence, incidence, mortality, and risk, from 1980 to present, and projected through 2040.[3] It is now updated regularly under the coordination by the Institute for Health Metrics and Evaluation (IHME), as are supporting estimates for sociodemographic and geospatial indicators. Each update is increasingly specific, including more detailed age categorization, causes, and "sub-national" (state/provincial or county-level) locations. The extent to which these estimates can inform public health policy has yet to be fully realized.

In previous work, an IHME team reported its development of a "Cost-Effectiveness Analysis with Microsimulation" (CEAM) tool, which has since been renamed "Vivarium."[7] Vivarium is an open-source, customizable, simulation model framework built on standard scientific Python libraries.[4] Designed to support generic simulations through either individual-based (microsimulation) or population-based (cohort-component) methods, Vivarium also has been enhanced with integrated tools to harness estimates from the GBD and expedite development of healthcare simulations.

Cost-Effectiveness Analysis (CEA) is a common and effective tool for healthcare policy planning, and microsimulation models like those run in Vivarium are a commonly used CEA technique.[6][2] However, the utility of current microsimulations is often constrained by three main limitations: scope, timeliness, and reproducibility. Models often limit the number or type of policy interventions under study, typically exploring only one or two similar strategies, or comparing a handful of strategies for which modeling approaches are essentially the same. There is often a considerable investment in the modeling pipeline, especially in data acquisition and management, but also in

software development, model processing, and validation. Developmental redundancy across the field contributes to wasted time and effort. As the CEAM development team discovered, open-source microsimulation models are often abandoned or lacking in sufficient documentation to warrant long-term adoption.[7] To address these issues, the microsimulation process must move toward faster, more efficient, and more transparent approaches.

The urgency of addressing these constraints became clear to IHME's simulation science team as they attempted to expand Vivarium to meet the needs of a current venture, which will be referred to in this paper as the "Simulated Child Health Interventions" (SimCHI) project. Its work seeks to inform seeks to inform optimized reduction of under-five childhood mortality in low- and middle-income settings, to meet the UN's Sustainable Development Goals.[12] The project's broad scope, requiring rapid investigation of many and variable intervention strategies applied across variable nations, necessitated a more efficient modeling workflow and the ability to create variable disease intervention packages. Presented here are the results of this developmental process in the context of the SimCHI project.

Significance   This research seeks to enhance the capabilities of Cost-Effectiveness Analyses toward greater complexity by improving process efficiency. To do so will increase the analytical bandwidth of the CEA modeling, leading to improved utility and accuracy of prediction.

2   Methods

2.1   Vivarium

The Vivarium core is designed to support general simulation modeling and can be applied to a wide spectrum of data types and research fields. Thus, input data from any source can be used: neither using GBD data nor even creating healthcare models is required to benefit from using the tool. However, this core has been expanded upon to include a regularly updated suite of healthcare-specific discrete-time Monte Carlo modeling tools, titled "Vivarium Public Health."[5] Using GBD estimates as the starting state, Vivarium Public Health runs simulations that adjust individual health status as defined by user-selected "model components", Python functions included in the Vivarium Public Health library.[1] Model components are customizable for each feature of the model, for example by the location, age, sex, and year. Uncertainty intervals are generated using draw-level inputs from the GBD, with additional parameter uncertainty facilitated by the common-random-numbers approach.[9]

2.2   Project Requirements

As of submission, more than fifty individual interventions have been flagged for potential development in the SimCHI project. These interventions were selected through expert opinion paired with a Systematic Review of Systematic Reviews. Each proposed intervention must first be investigated for individual viability (i.e. "can a defensible model be produced from the available data?").[2] Viable strategies must then be modeled alone before being added to the pool of potential strategies to be bundled in "packages", variable and interchangeable groupings of interventions.

---

[1] Users can also create bespoke components for their own domain- or simulation-specific needs, although this task typically requires advanced software development experience.

[2] This task is still in progress.

2.3 Developmental Goals

To successfully manage the complexity of SimCHI, the team set the following process improvement goals:

1: Develop software to flexibly specify modular intervention packages;

2: Develop workflow to support fast turn-around and parallel development of new models and methods;

3: Create a documentation strategy that will enable reproduction of our final results using only the resources published online.

3 Results

3.1 SimCHI Project

The earliest SimCHI models are health interventions for the improvement of childhood malnutrition and related mortality. GBD data are used as input values for estimates such as predicted population, vaccine coverage, measures of risk exposure, and malnutrition disease burden. Inputs may also be supplemented by data from randomized controlled trials, for example to establish intervention effectiveness.

Thus far the team have completed individual models for shigella vaccination, a breastfeeding promotion campaign, dietary fortification and dietary supplementation. Meanwhile, the team continues to produce individual models as it develops its first packages using the malnutrition interventions. The results of these models will be discussed in future publications.

3.2 Modeling Workflow

The team developed an iterative and explicit "development workflow," containing a clear separation of tasks by skillset and encouraging specialization. This workflow involves three major stages and eight phases, each with templated phase output documentation (See Figure 1 and Figure 2). Using this workflow, the team observed a significant increase in productivity.
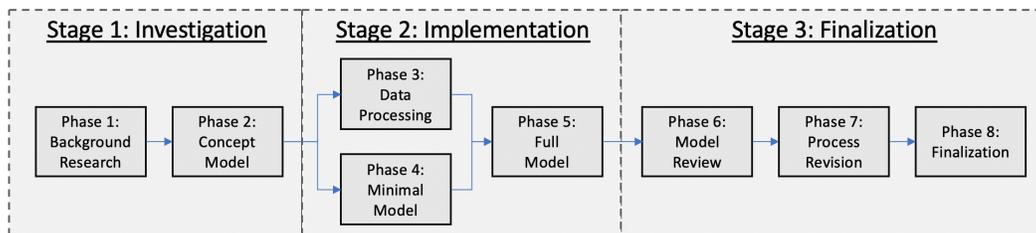

Figure 1: Model Development Workflow

Each intervention model begins with a preliminary Investigation stage (Stage 1) led by the researcher. First, a fast survey of the intervention literature is performed in the form of a "Systematic Search" to quickly scan and track reliable sources to determine if there is enough evidence for a defensible model. At this stage the modeler also begins an Intervention Report, a template that guides the modeling process and serves as an active record of major modeling decisions. Ideally the report is updated at the end of each phase of development. Then, modeling feasibility is assessed, informed by input from domain experts working on the GBD. When a sound approach is determined, model requirements are tracked on a formalized Concept Model document. Once the

researcher and engineer have agreed on their shared understanding of the requirements as explicitly stated in the Concept Model, the developer begins leading the Implementation stage.

Implementation (Stage 2) begins with data extraction and analysis, both of GBD estimates and of supplementary input data. This process may be as simple as recording a single number from one randomized controlled trial, or it might be as complex as running several Systematic Reviews (informed by the Systematic Search) that lead to complex regressions against GBD covariates. Once input data are processed and vetted, the model undergoes two rounds of component development: a "minimal model" and a "full model". The minimal model is the easiest-to-produce, defensible version of the intervention. Complex interactions may be simplified, or specific trends may be generalized, providing the researcher with intermediate results on which they can test their model validation strategies as additional model components are constructed. For example, the minimal model may be the equivalent of a 'magic wand' that suddenly increases vaccine coverage for all individuals, whereas the full model would use antenatal care coverage, quality, and costing components to create a more realistic scenario.

The last stage, Finalization (Stage 3), is a joint effort focused on resolution and incremental process improvement. Outstanding tasks are assessed, successes and failures are discussed, and specific plans are made to address both. Lastly, each model is attached to a shareable "Data Artifact" and published online in a public code repository (containing the model code and pinned dependencies), allowing individual models to be externally cloned and analyzed, and further enabling collaboration and transparency.[5]

4   Discussion

4.1   Using GBD Inputs

There are two notable constraints when building these GBD-facilitated models. First, all models must be tailored to meet at least the cause-level categories outlined by the GBD. This can present minor challenges when, for example, interventions address outcomes that are included in multiple GBD causes or are estimated as one fraction of a single parent cause. Second, because of the scale of the study, estimates are as yet only available for national and a select number of sub-national populations. However, the GBD study is now updated annually and with increasing precision. Each iteration adds more stratification of identifiers (like GBD-causes and estimated locations), working toward the 5x5 kilometer estimates that are the goal of the "Local Burden of Disease."[8] Hence, this second hurdle may become a dwindling issue in the future. Moreover, because of the iterative updates, models built using Vivarium's GBD-integrated tools can be updated to include that increased precision with minimal work.

There are noteworthy benefits to a GBD-integrated model. The first comes at the validation stage. Among other techniques, the team are now using a "temporal knockout" test, in which disease models are run using a temporally truncated version of input data (for example, an input with data for only years 1980 though 2010). Results for the temporal knockout can then be compared to actual GBD estimates, and this works for either retrospective or prospective models. Such a test is often not possible because of temporal limitations in the input data. Second, models are easily translatable between locations. A model for a location with a wide margin of error can also be tested on a location with more precise estimates. This eliminates data quality as a potential problem and provides more information to the modeler. Again, this type of validation is unavailable to models that rely on input data of a smaller spatiotemporal scope. Finally, the greatest advantage of a GBD-assisted model is that the vast majority of the input data acquisition, management, and

	Researcher Tasks	Engineer Tasks	Output
Phase 1: Background Research	* Systematic Search * Brainstorm model approaches * Consult with GBD domain experts and team lead	* Develop tools in support of background research and input data analytics	**Researcher** * Systematic Search Document * Start Intervention Report **Engineer** * Completed tools (as applicable)
Phase 2: Concept Model	* Attend Concept Model Development Meeting * Plan validation for input data	* Attend Concept Model Development Meeting	**Researcher** * Concept Model Document * Updated Intervention Report
Phase 3: Data Processing	* Extract and format data for effect size, existing coverage, and cost * Validate input data processing * Develop validations and visualizations	* Support data extraction and generate Data Artifact	**Researcher** * Systematic Review Document(s) * Processed Supplemental Input(s) * Updated Intervention Report **Engineer** * Data Artifact
Phase 4: Minimal Model	* Launch test models	* Develop and test minimal model per the Concept Model	**Researcher** * Updated Intervention Report **Engineer** * Functioning Minimal Model
Phase 5: Full Model	* Finalize validations for output data * Run small scale (10-100 draws) validation model	* Replace test components with data-driven components * Develop custom Observer components * Add health intervention data to Data Artifact	**Researcher** * Validated Full Model * Updated Intervention Report **Engineer** * Functioning Full Model * Updated Data Artifact
Phase 6: Model Review	* Present results to team for final validation * Run final models * Complete model visualizations	*Incorporate final model corrections * Hold final code review for the model	**Researcher** * Final Intervention Report * Output Visualizations
Phase 7: Process Revision	* Present final results to client, as needed * Participate in Process Improvement Discussion	* Plan future tools to support input data analytics (see Phase 1) * Participate in Process Improvement Discussion	**Researcher and Engineer** * Final Documentation * Shared Plan for Process Improvement
Phase 8: Finalization		* Generalize model features for future use * Publish Open-Source Replication Archive	**Engineer** * Archived Model

Figure 2: Model Development Phases, Outputs, and Responsibilities

validation is handled by other teams, specialists working on the GBD study. The Simulation Science team itself must do only minimal data seeking and data cleaning. This is one of the most major time sinks in any data modeling project, and reducing the requirement allows the team to focus on model production. Moreover, relying on domain experts to do the proverbial heavy lifting provides additional assurance of model accuracy.

4.2   Establishing the Workflow

As has been observed in many other industries, the team found that clear communication with clearly defined responsibilities significantly increased efficiency after the prototyping stage. The greatest boon observed was increased project parallelizebility. By specializing in clearly defined tasks and by actively tracking progress with clear templates, the tasks themselves have become interchangeable among models. This means projects can be stopped, switched, and re-started with increased ease. It also allows tasks and projects to easily change ownership as needed, and reduces training overhead (since documentation can be used for training). Furthermore, specialization leads to efficiency by reducing time lost in the mental transition between problem types. However, the team did observe that the existence of clear references (documentation and templates) can reduce required mental transition time, indicating that the benefits to using this type of active-documentation workflow are still present for teams where one person might assume the roles of both researcher and software developer.

4.3   Facilitating Transparency

Clear documentation is a specific priority for the team in order to foster reproducible research[10] and to support our institutional commitment to comply with the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER).[11] Viviarum installation and setup instructions, as well as preliminary technical documentation are already available online.[5] Additionally, the team is developing a series of tutorials and quickstart materials to be published with the Vivarium documentation alongside our phase output templates within the next few months.

As is common for projects of this type, intermediate adjustments to staffing and scope led to a prioritization of results at the expense of documenting our research processes. However, with the recent workflow changes the team are now actively tracking model progress through phase-specific templates, and we believe we can iteratively update those templates such that they are still useful internally while providing legible-enough information to inform an outside audience (given the context of the GBD). Moreover, the team's Software Engineers now hold bi-weekly documentation working sessions to continue updating and polishing the technical materials for legibility. Through this method we intend to reduce our documentation requirements and increase the likelihood of completing external documentation at the same time as the model.

4.4   Looking Ahead

The team has identified new intermediate goals in continued pursuit of greater modeling throughput. The number one priority is, of course, to complete the models for currently identified interventions. Among other goals is a collaboration to develop systematic approaches for researching and managing complex disease interactions: it can be difficult in epidemiological modeling to find studies that adequately address time-lagged or combinatory effects among the interventions in question, at least such that they can be used in our approach, so this may well be the next greatest hurdle to increasing our modeling capacity. The team is also working on methods to enable rapid comparison of differing intervention packages. Lastly, the team is investigating new operational and computational strategies to manage interventions and packages as they accumulate at scale.

5   Conclusion

IHME's Simulation Science team continued the development of its Vivarium tool in pursuit of the SimCHI project, an effort to reduce the burden of childhood mortality through enhanced cost-effectiveness modeling. With the long-term goal of comparing many modular health intervention

packages, the team set three intermediate goals toward that end: a method for implementing those modular interventions; a workflow facilitating the rapid development of those interventions; and a method for providing transparency through documentation. These goals are at or near completion as the team looks ahead to manage intervention packages at scale.

6 Acknowledgements

References

[1] Institute for Health Metrics and Evaluation. Global Burden of Disease. Accessed 2019, Feb 1. http://www.healthdata.org/gbd.

[2] Siebert U, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force – 3. Value Health. 2012, Sep-Oct; 15(6):812-20. doi: 10.1016/j.jval.2012.06.014.

[3] Institute for Health Metrics and Evaluation. Future Health Scenarios. Accessed 2019, Feb 1. http://www.healthdata.org/future-health-scenarios.

[4] Collins J, et al. Vivarium V0.8.15. Zenodo. 2019, Jan 14. doi: 10.5281/zenodo.2539174.

[5] Collins J, et al. Vivarium Public Health V0.8.13. Zenodo. 2019, Jan 14. doi:10.5281/zenodo.2539177.

[6] Neumann P J, Sanders G D, Russell L B, Siegel J, Ganiats, T G. Cost-Effectiveness in Health and Medicine. Oxford University Press. 2016.

[7] Sorensen R, Flaxman A, Deason A, Mumford J E, Eldrenkamp E, Moses M, Weaver M. Microsimulation models for cost-effectiveness analysis: a review and introduction to CEAM. SummerSim-SCSC: Proceedings of the Summer Simulation Multi-Conference, 2017, Jul 9.

[8] Institute for Health Metrics and Evaluation. Local Burden of Disease. Accessed 2019, Feb 1. http://www.healthdata.org/lbd.

[9] Flaxman A, Deason A, Dolgert A, Mumford J, Sorensen R, Eldrenkamp E, Vos T, Foreman K, Mokdad A, Weaver M. Untangling Uncertainty with Common Random Numbers: A Simulation Study. SummerSim-SCSC: Proceedings of the Summer Simulation Multi-Conference. 2017, Jul 9.

[10] Sandve G K, Nekrutenko Anton, Taylor J, Hovig E. Ten Simple Rules for Reproducible Computational Research. PLoS Comput Biol. 2013, Oct; 9(10): e1003285.

[11] Institute for Health Metrics and Evaluation. GATHER – A new age for global health metrics. 2016, Jun 29. http://www.healthdata.org/acting-data/gather-new-age-global-health-metrics.

[12] United Nations Children's Fund (UNICEF). Child survival and the SDGs. 2017, Oct. https://data.unicef.org/topic/child-survival/child-survival-sdgs.