Simulation-Based Inference for Global Health Decisions

Christian Schroeder de Witt 1 Bradley Gram-Hansen 1 Nantas Nardelli 1
Andrew Gambardella 1 Rob Zinkov 1 Puneet Dokania 1 N. Siddharth 1
Ana Belen Espinosa-Gonzalez 2 Ara Darzi 2 Philip Torr 1 Atılım Güneş Baydin 1

Abstract
The COVID-19 pandemic has highlighted the importance of in-silico epidemiological modelling in predicting the dynamics of infectious diseases to inform health policy and decision makers about suitable prevention and containment strategies. Work in this setting involves solving challenging inference and control problems in individual-based models of ever increasing complexity. Here we discuss recent breakthroughs in machine learning, specifically in simulation-based inference, and explore its potential as a novel venue for model calibration to support the design and evaluation of public health interventions. To further stimulate research, we are developing software interfaces that turn two cornerstone COVID-19 and malaria epidemiology models (CovidSim1 and OpenMalaria2) into probabilistic programs, enabling efficient interpretable Bayesian inference within those simulators.

1 Introduction
Machine learning has a growing role in increasing health service access and efficiency, particularly in resource-constrained settings, making it a valuable tool for the global health community [39, 54]. Moreover, the COVID-19 pandemic [55] has underlined the importance of epidemiological modelling and computer simulation in informing the design and implementation of public health interventions at an unprecedented scale [18]. For many endemic diseases (e.g., malaria), in-silico optimisation of multi-modal intervention portfolios—from mass vaccination to bed nets—is well established [47]. Analogous modelling for COVID-19 interventions, including social distancing [20], is mostly unexplored, yet subject to intense public interest [32].

The adoption of health informatics in worldwide health systems (e.g., OpenMRS [33], mHealth [1]) enables access to abundant patient-level and aggregated health data [54]. This is fomenting the development of comprehensive modelling and simulation to support the design of health interventions and policies, and to guide decision-making in a variety of health system domains [22, 49]. For example, simulations have provided valuable insight to deal with public health problems such as tobacco consumption in New Zealand [50], and diabetes and obesity in the US [58]. They have been used to explore policy options such as those in maternal and antenatal care in Uganda [44], and applied to evaluate health reform scenarios such as predicting changes in access to primary care services in Portugal [21]. Their applicability in informing the design of cancer screening programmes has been also discussed [42, 23]. Recently, simulations have informed the response to the COVID-19 outbreak [19].

The process of informing health interventions and policies through simulations generally involves two steps:

Model calibration The extent to which a simulator can reliably inform real-world prediction and planning is bounded by both model discrepancy [13] and how well the model has been calibrated to empirical data [3].

Optimising decision-making Identifying optimal multimodal intervention strategies and corresponding risks and uncertainties requires searching through potentially vast parameter spaces, which, due to the computational cost of running large simulators (e.g., in some epidemiological studies), usually cannot be exhaustively evaluated [46].

Despite their fundamental importance, model discrepancy and calibration of public-health simulators are frequently only informally addressed, or left undocumented [48, 40]. This may be partially explained by the fact that, while numerous methods for formal sensitivity and uncertainty analysis exist [28], they in general do not scale to complex simulators with more than a few dozen parameters [38]. Similarly, evidence-based decision-making is usually optimised by comparing outcomes on a small number of hand-crafted scenarios and intervention strategies [46].

2 Epidemiology simulations and inference
Among the simplest mathematical epidemiology models are deterministic compartmental models that partition individu-
Simulations in a population based on different stages of the disease\(^3\) [29, 2, 11]. Advances in model construction, computing power, and novel insights into medical and socio-economic aspects have since stimulated the introduction of stochastic individual-based models \(^4\) [41] to public health applications. These relatively complex and highly-parametrised models are implemented as simulator software that allow studying the global effects of self-organization and emergent properties arising from individual interactions at the local level.

In general, inference in individual-based simulators is usually doubly intractable, as both simulator likelihood and evidence cannot be evaluated efficiently. Likelihood-free methods, including approximate Bayesian computation (ABC) [9] have been proposed [3, 16], but suffer from exponential scaling of inference with data dimension, requiring domain experts to define low-dimensional summary statistics, which ultimately determine quality of inference.

Recent advances in machine learning have led to a new family of promising approaches to simulation-based inference [see 15, for an overview]. In particular, we argue that probabilistic programming [see 53, for an introduction] has a unique potential to standardise and automate model calibration and decision-making in individual-based simulators.

Probabilistic programming allows one to express probability models using computer code and perform statistical inference over the inputs and latent variables of the program, conditioned on data observations (or constraints). This is achieved by using special-purpose probabilistic programming languages (PPLs) [24, 43, 14, 51, 10], which augment a host language with features to express probabilities and Bayesian conditioning. PPLs separate model specification from inference, allowing flexible selection of appropriate inference engines (e.g., importance sampling (IS) [17], Markov-chain Monte Carlo (MCMC) [12]).

Recent work made it possible to use pre-existing stochastic simulators as probabilistic programs, with minimal code modification to capture and redirect random number draws [7], scaling up to very large simulators [8], and particularly relevant to this work, with application to individual-based epidemiology simulators [26]. Within such a framework, one could, for instance, condition on desired health outcomes (e.g., ICU capacity not being exceeded in a pandemic), and derive detailed posterior distributions over all interactions defined by the simulator [56], providing insights on interventions effecting a desired outcome—with proper uncertainty quantification at all stages.

To further enhance the applicability of simulation-based inference in this domain we highlight several opportunities for further method development. Automated amortisation by surrogate methods [25, 36, 34], which aim to automatically identify and replace compute-intensive parts of a simulator through less expensive emulators, could be guided by the causal structure inherent to a simulator (Figure 1), such as many repeated, structurally identical stochastic time steps or multi-agent interactions that might be amenable to field approximations [57]. In addition, pre-existing simulators could be turned into differentiable programs [6] through automated source-to-source transformations, thus allowing for the use of gradient-based optimisation and inference methods, including Hamiltonian Monte Carlo [37]. Last but not least, the unified interface specification afforded by probabilistic programming could allow simulators to also become amenable to other techniques from simulation-based inference and control, including dynamic programming and reinforcement learning [31, 52, 27].

To foster the development of a new standardised approach to model calibration and evidence-based decision-making in public health, we are working on instrumenting the existing CovidSim [19] and OpenMalaria [45] simulators with a probabilistic programming interface through the PyProb library.\(^5\) We will publicly release our code to provide out-of-the-box probabilistic programming inference over public health scenarios of interest in these two domains.

We expect the mentioned techniques to play a role in dealing with communicable (infectious) diseases, which already entailed a significant burden for health systems in developing countries [35] before the worldwide impact of COVID-19. However, they can also be applicable to non-communicable diseases, such as diabetes and cancer, which are recognised major causes of morbidity and mortality worldwide [4, 30]. This will add to the already identified potential of machine learning in health policy [5] and improving health access, emphasising its value for global health in the efforts to achieve universal health coverage and sustainable development goals [39, 54].

\(^1\)E.g., susceptible–infectious–recovered or SIR

\(^2\)Also referred to as agent-based model or multi-agent system.

\(^3\)https://github.com/pyprob/pyprob
References

[1] Sasan Adibi. *Mobile Health: A Technology Road Map*, volume 5. Springer, 2015.

[2] Linda JS Allen. Some discrete-time SI, SIR, and SIS epidemic models. *Mathematical Biosciences*, 124(1): 83–105, 1994.

[3] Ioannis Andrianakis, Ian R. Vernon, Nicky McCreesh, Trevelyan J. McKinley, Jeremy E. Oakley, Rebecca N. Nsubuga, Michael Goldstein, and Richard G. White. Bayesian History Matching of Complex Infectious Disease Models Using Emulation: A Tutorial and a Case Study on HIV in Uganda. *PLOS Computational Biology*, 11(1):e1003968, January 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1003968. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003968.

[4] John P Ansah, Amina Mahmood Islam, Victoria Koh, Vanthy Ly, Hero Kol, David B Matchar, Chhun Loun, and Mondol Loun. Systems modelling as an approach for understanding and building consensus on non-communicable diseases (NCD) management in Cambodia. *BMC Health Services Research*, 19(1):1–10, 2019.

[5] Hutan Ashrafian and Ara Darzi. Transforming health policy through machine learning. *PLoS Medicine*, 15(11), 2018.

[6] Atılım Güneş Baydin, Barak A. Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research (JMLR)*, 18(153):1–43, 2018. URL http://jmlr.org/papers/v18/17-468.html.

[7] Atılım Güneş Baydin, Lukas Heinrich, Wahid Bhimji, Lei Shao, Saeid Naderiparizi, Andreas Munk, Jialin Liu, Bradley Gram-Hansen, Gilles Louppe, Lawrence Meadows, Philip Torr, Victor Lee, Prabhat, Kyle Cranmer, and Frank Wood. Efficient probabilistic inference in the quest for physics beyond the standard model. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, 2019. URL http://papers.nips.cc/paper/8785-efficient-probabilistic-inference-in-the-quest-for-physics-beyond-the-standard-model.pdf.

[8] Atılım Güneş Baydin, Lei Shao, Wahid Bhimji, Lukas Heinrich, Lawrence Meadows, Jialin Liu, Andreas Munk, Saeid Naderiparizi, Bradly Gram-Hansen, Gilles Louppe, Mingfei Ma, Xiaohui Zhao, Philip Torr, Victor Lee, Kyle Cranmer, Prabhat, and Frank Wood. Etalumis: Bringing Probabilistic Programming to Scientific Simulators at Scale. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–24, November 2019. doi: 10.1145/3295500.3356180. URL http://arxiv.org/abs/1907.03382. arXiv: 1907.03382.

[9] Mark A Beaumont, Wenyang Zhang, and David J Balding. Approximate Bayesian computation in population genetics. *Genetics*, 162(4):2025–2035, 2002.

[10] Eli Bingham, Jonathan P Chen, Martin Jankowiak, Fritz Obermeyer, Neeraj Pradhan, Theofanis Karaletsos, Rohit Singh, Paul Szerlip, Paul Horsfall, and Noah D Goodman. Pyro: Deep universal probabilistic programming. *The Journal of Machine Learning Research*, 20(1):973–978, 2019.

[11] Fred Brauer. Compartmental models in epidemiology. In *Mathematical Epidemiology*, pages 19–79. Springer, 2008.

[12] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. CRC press, 2011.

[13] Jenny Brynjarsdóttir and Anthony O’Hagan. Learning about physical parameters: The importance of model discrepancy. *Inverse Problems*, 30(11):114007, 2014.

[14] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 2017.

[15] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The frontier of simulation-based inference. *arXiv:1911.01429 [cs, stat]*, April 2020. URL http://arxiv.org/abs/1911.01429. arXiv: 1911.01429.

[16] Katalin Csilléry, Michael GB Blum, Oscar E Gaggiotti, and Olivier François. Approximate Bayesian computation (ABC) in practice. *Trends in Ecology & Evolution*, 25(7):410–418, 2010.

[17] Arnaud Doucet and Adam M Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12(656-704):3, 2009.

[18] S. Eubank, I. Eckstrand, B. Lewis, S. Venkataraman, M. Marathe, and C. L. Barrett. Commentary on Ferguson, et al., “Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand”. *Bulletin of Mathematical Biology*, 82(4), 2020. ISSN
Neil Ferguson, Daniel Laydon, Gemma Nedjati Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, ZULMA Cucunuba Perez, Gina Cuomo-Dannenburg, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. 2020.

Neil M. Ferguson, Derek A. T. Cummings, Christophe Fraser, James C. Cajka, Philip C. Cooley, and Donald S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, July 2006. ISSN 1476-4687. doi: 10.1038/nature04795. URL https://www.nature.com/articles/nature04795.

André S Fialho, Mónica D Oliveira, and Armando B Sá. Using discrete event simulation to compare the performance of family health unit and primary health care centre organizational models in Portugal. *BMC Health Services Research*, 11(1):274, 2011.

David Fone, Sandra Hollinghurst, Mark Temple, Alison Round, Nathan Lester, Alison Weightman, Katherine Roberts, Edward Coyle, Gwyn Bevan, and Stephen Palmer. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health*, 25(4):325–335, 12 2003. ISSN 1741-3842. doi: 10.1093/pubmed/fdg075. URL https://doi.org/10.1093/pubmed/fdg075.

Abraham M Getaneh, Eveline Heijnsdijk, and Harry de Koning. The role of modelling in the policy decision making process for cancer screening: example of prostate specific antigen screening. *Public Health Research & Practice*, 29(2), 2019.

Noah Goodman, Vikash Mansinghka, Daniel M Roy, Keith Bonawitz, and Joshua B Tenenbaum. Church: a language for generative models. *arXiv preprint arXiv:1805.00909 [cs, stat]*, May 2018. URL http://arxiv.org/abs/1805.00909. arXiv: 1805.00909.

Elisabeth Mahase. Covid-19: public health experts demand evidence behind UK’s short self-isolation advice. *BMJ*, 369, 2020. doi: 10.1136/bmj.m1574. URL https://www.bmj.com/content/369/bmj.m1574.

Burke W Mamlin, Paul G Biondich, Ben A Wolfe, Hamish Fraser, Darius Jazayeri, Christian Allen, Justin Miranda, and William M Tierney. Cooking up an open source EMR for developing countries: OpenMRS–a recipe for successful collaboration. In *AMIA Annual Symposium Proceedings*, volume 2006, page 529. American Medical Informatics Association, 2006.

Andreas Munk, Adam Šćibor, Atılım Güneş Baydín, Andrew Stewart, Goran Fernlund, Anoush Pour-sartip, and Frank Wood. Deep Probabilistic Surro-
gate Networks for Universal Simulator Approximation. arXiv:1910.11950 [cs, stat], October 2019. URL http://arxiv.org/abs/1910.11950. arXiv: 1910.11950.

[35] Christopher JL Murray, Katrina F Ortblad, Caterina Guinovart, Stephen S Lim, Timothy M Wolock, D Allen Roberts, Emily A Dansereau, Nicholas Graetz, Ryan M Barber, Jonathan C Brown, et al. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the global burden of disease study 2013. The Lancet, 384(9947):1005–1070, 2014.

[36] Saeid Naderiparizi, Adam ´Scibior, Andreas Munk, Mehrdad Ghadiri, Atılım Güne¸s Baydin, Bradley Gram-Hansen, Christian Schroeder de Witt de Witt, Robert Zinkov, Philip H.S. Torr, Tom Rainforth, Yee Whye Teh, and Frank Wood. Amortized rejection sampling in universal probabilistic programming. arXiv preprint arXiv:1910.09056, 2019.

[37] Radford M. Neal. MCMC using Hamiltonian dynamics. arXiv:1206.1901 [physics, stat], June 2012. URL http://arxiv.org/abs/1206.1901. arXiv: 1206.1901.

[38] Jeremy E Oakley and Anthony O’Hagan. Probabilistic sensitivity analysis of complex models: a Bayesian approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66(3):751–769, 2004.

[39] Trishan Panch, Peter Szolovits, and Rifat Atun. Artificial intelligence, machine learning and health systems. Journal of Global Health, 8(2), 2018.

[40] Narat Punyacharoensin, William John Edmunds, Daniela De Angelis, and Richard Guy White. Mathematical models for the study of HIV spread and control amongst men who have sex with men. European Journal of Epidemiology, 26(9):695, 2011.

[41] Steven F Railsback and Volker Grimm. Agent-Based and Individual-Based Modeling: A Practical Introduction. Princeton University Press, 2019.

[42] Guglielmo Ronco and Iacopo Baussano. Causal system modelling of cervical cancer screening. The Lancet Public Health, 2(2):e61–e62, 2017.

[43] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. Probabilistic programming in Python using PyMC3. PeerJ Computer Science, 2:e55, 2016.

[44] Agnes Rwashana Semwanga, Sarah Nakululwa, and Taghreed Adam. Applying a system dynamics modelling approach to explore policy options for improving neonatal health in Uganda. Health Research Policy and Systems, 14(1):35, 2016.

[45] Thomas Smith, Gerry F Killeen, Nicolas Maire, Amanda Ross, Louis Molineaux, Fabrizio Tediosi, Guy Hutton, Jürg Utzinger, Klaus Dietz, and Marcel Tanner. Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of plasmodium falciparum malaria: Overview. The American Journal of Tropical Medicine and Hygiene, 75(2_suppl):1–10, 2006.

[46] Thomas Smith, Amanda Ross, Nicolas Maire, Nakul Chitnis, Alain Studer, Diggory Hardy, Alan Brooks, Melissa Penny, and Marcel Tanner. Ensemble Modeling of the Likely Public Health Impact of a Pre-Erythrocytic Malaria Vaccine. PLOS Medicine, 9(1):e1001157, January 2012. ISSN 1549-1676. doi: 10.1371/journal.pmed.1001157. URL https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001157.

[47] Thomas A. Smith, Nakul Chitnis, Melissa Penny, and Marcel Tanner. Malaria Modeling in the Era of Eradication. Cold Spring Harbor Perspectives in Medicine, 7(4):a025460, January 2017. ISSN , 2157-1422. doi: 10.1101/cshperspect.a025460. URL http://perspectivesinmedicine.cshlp.org/content/7/4/a025460. Publisher: Cold Spring Harbor Laboratory Press.

[48] Natasha K Stout, Amy B Knudsen, Chung Yin Kong, Pamela M McMahon, and G Scott Gazelle. Calibration methods used in cancer simulation models and suggested reporting guidelines. Pharmacoeconomics, 27(7):533–545, 2009.

[49] R Chad Swanson, Adriano Cattaneo, Elizabeth Bradley, Somsak Chunharas, Rifat Atun, Kaja M Abbas, Korina Katsaliaki, Navonil Mustafee, Benjamin Mason Meier, and Allan Best. Rethinking health systems strengthening: key systems thinking tools and strategies for transformational change. Health Policy and Planning, 27(suppl_4):iv54–iv61, 2012.

[50] Martin I Tobias, Robert Y Cavana, and Ashley Bloomfield. Application of a system dynamics model to inform investment in smoking cessation services in New Zealand. American Journal of Public Health, 100(7):1274–1281, 2010.

[51] Dustin Tran, Matthew D. Hoffman, Dave Moore, Christopher Suter, Sriravas Vasudevan, Alexey Radul, Matthew Johnson, and Rif A. Saurous. Simple, distributed, and accelerated probabilistic programming. In Neural Information Processing Systems, 2018.

[52] Jan-Willem van de Meent, Brooks Paige, David Tolpin, and Frank Wood. Black-Box Policy Search with Probabilistic Programs. Proceedings of the 19th Interna-
Simulation-Based Inference for Global Health Decisions

tional Conference on Artificial Intelligence and Statistics, page 1195–1204, 2016.

[53] Jan-Willem van de Meent, Brooks Paige, Hongseok Yang, and Frank Wood. An introduction to probabilistic programming. arXiv preprint arXiv:1809.10756, 2018.

[54] Brian Wahl, Aline Cossy-Gantner, Stefan Germain, and Nina R Schwalbe. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Global Health, 3(4), 2018. doi: 10.1136/bmjgh-2018-000798. URL https://gh.bmj.com/content/3/4/e000798.

[55] WHO. Coronavirus, 2020. URL https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Library Catalog: www.who.int.

[56] Frank Wood, Andrew Warrington, Saeid Naderiparizi, Christian Weilbach, Vaden Masrani, William Harvey, Adam Scibor, Boyan Beronov, and Ali Nasseri. Planning as Inference in Epidemiological Models. arXiv:2003.13221 [cs, q-bio, stat], April 2020. URL http://arxiv.org/abs/2003.13221. arXiv: 2003.13221.

[57] Yaodong Yang, Rui Luo, Minne Li, Ming Zhou, Weinan Zhang, and Jun Wang. Mean Field Multi-Agent Reinforcement Learning. arXiv:1802.05438 [cs], July 2018. URL http://arxiv.org/abs/1802.05438. arXiv: 1802.05438.

[58] Donglan Zhang, Philippe J Giabbanelli, Onyebuchi A Arah, and Frederick J Zimmerman. Impact of different policies on unhealthy dietary behaviors in an urban adult population: an agent-based simulation model. American Journal of Public Health, 104(7): 1217–1222, 2014.