 doi:10.20944/preprints202206.0137.v1

Communication

FAIR Sharing of Reproducible Models of Epidemic and Pandemic Forecast

Kausthubh Ramachandran[1§], Matthias König[2§], Martin Scharm[3], Tung V. N. Nguyen[1], Henning Hermjakob[1], Dagmar Waltemath[4*] and Rahuman S Malik Sheriff [1*]

[1] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, UK
[2] Institute for Theoretical Biology, Humboldt-University Berlin, Berlin, Germany
[3] Department of Systems Biology and Bioinformatics, University of Rostock, 18051 Rostock, Germany
[4] Department of Medical Informatics, University Medicine Greifswald, 17475 Greifswald, Germany
[§] shared first authors
*Corresponding author: sheriff@ebi.ac.uk (R.S.M.S.); dagmar.waltemath@uni-greifswald.de (D.W.)

Abstract: A major challenge for the dissemination, replication, and reuse of epidemiological forecasting studies during COVID-19 pandemics is the lack of clear guidelines and platforms to exchange models in a Findable, Accessible, Interoperable, and Reusable (FAIR) manner, facilitating reproducibility of research outcomes. During the beginning of pandemics, models were developed in diverse tools that were not interoperable, opaque without traceability and semantics, and scattered across various platforms - making them hard to locate, infer and reuse. In this work, we demonstrate that implementing the standards developed by the systems biology community to encode and share COVID-19 epidemiological models can serve as a roadmap to implement models as a tool in medical informatics, in general. As a proof-of-concept, we encoded and shared 24 epidemiological models using the standard format for model exchange in systems biology, annotated them with cross-references to data resources, packed up all associated files in COMBINE archives for easy sharing, and finally, disseminated the models through BioModels repository to significantly enhance their reproducibility and repurposing potential. We recommend the use of systems biology standards to encode and share models of epidemic and pandemic forecasts to improve their findability, accessibility, interoperability, reusability, and reproducibility.

Keywords: FAIR; epidemiology; models; pandemic forecast; SIR modelling; standards

1. Introduction

The early stages of the COVID-19 pandemic have proved that there is a critical need to improve sharing of defensible and reusable models of disease transmission. When the pandemic started in 2019, the shared disease transmission simulations were opaque and lacked traceability to model structure, semantics, and connections to the source knowledge, resulting in their failure to provide the contexts and insights required to support decisions. It was almost impossible to adapt the epidemiological models in a timely manner due to the difficulty in extracting information from model documents (1). Though mathematical modeling of infectious disease transmission using ordinary differential equations (ODEs) has been central to epidemiology for decades, reporting guidelines on epidemic forecasting did not exist (2).

In 1927, Kermack and McKendrick first published the Susceptible-Infected-Recovered (SIR) model to forecast the growth of epidemics (3). In SIR models, the population is usually divided into mutually exclusive compartments and the spread of disease between compartments happens via interaction between them. The SIR model and its derivative versions are used to estimate key parameters, forecast trends in the variation of different variables such as proportion of susceptible population, daily infections, mortality, and test the validity of hypotheses using available data. Disease transmission models for

 doi:10.20944/preprints202206.0137.v1

SARS-CoV-2 are used by governments to implement non-pharmaceutical interventions like lockdowns across the globe (4). More specific transmission models for SARS-CoV-2 variants including the most contagious Delta and Omicron and their transition stages were also developed.

Ordinary differential equation (ODE) systems of epidemiological models representing disease transmission are typically encoded in a programming language to enable efficient numerical simulation, analysis, and automated visualizations of epidemic forecasts. Since 2020, the number of models simulating transmission dynamics of COVID-19 has increased tremendously. This has prompted researchers to put out calls for transparency in COVID-19 epidemiological research (5–7). It should be expected that published models and their results can be reproduced by other scientists with reasonable effort and thus can be trusted. In our recent study (8), we systematically analyzed 455 ODE models from published research across various fields of biomedical science. We found that the simulation results of nearly half of the published models could not be reproduced with the information provided in the associated publication including supplementary material. A smaller study for quantitative systems pharmacology models could execute only four of 12 models (9).

One major issue we exposed from our analysis is that model codes are hard to locate. They are usually spread across different resources like GitHub or personal websites, and the majority of models are not available on a common, publicly accessible repository. Another issue is poor interoperability between different modeling languages and tools. Model codes, when publicly available, are represented in different programming languages and software with diverse coding formats, and they often contain ambiguous annotations, versions, and comments which makes repurposing more difficult. Furthermore, misprints, typographical errors, missing or incomplete information, mismatch between the code and the description in the publication create more obstacles to reproduce, reuse, and repurpose models (8).

Any one or a combination of the listed roadblocks usually results in the scenario where a modeler is forced to spend a significant amount of time in retrieving and mining the model code, verifying code reproducibility, repurposing the code to work in the preferred programming environment, and then reusing the model as needed. If the above issues result in irreproducible models, researchers have to trust what is reported in the publication, and repurposing becomes impossible. To reduce costs and effort as well as to respect good scientific practices, it is essential to create and share epidemiological models, as any other simulation models, that are both reliable and reproducible to maximize utility and ensure widespread use. Particularly in a fast-evolving epidemic situation the time needed to reproduce a published result is a critical factor.

Several of these issues were recognized as roadblocks to reproducibility in epidemiological research as early as 2006 (10). Since then, the epidemiology community has taken significant steps to improve the transparency and reproducibility of their studies. However, unlike clinical data, epidemiological data were not shared on global platforms such as vivli.org (11). There have also been calls to post the code used to analyze the data publicly along with the data and results obtained from running the code (12).

The rapid increase in the number of published epidemiological forecast models during COVID-19 pandemics highlighted the problems that arose due to the lack of a consistent set of reporting guidelines. Hence, the EPIFORGE checklist (2) was developed to provide guidelines to report manuscripts describing models of epidemiological forecast. EPIFORGE focuses on clarity in reporting the context and assumptions of models. But these steps are not sufficient to ensure full mathematical reproducibility as they do not address the issues with model codes including poor interoperability across platforms, and semantic ambiguity. Hence, there is a need for additional measures to increase the reproducibility and ultimately the repurposing of epidemiological models.

The field of systems biology has faced similar issues and has spent years addressing them. In this article, we outline the measures adopted by the systems biology community to improve model reproducibility and reusability. Furthermore, we demonstrate that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 June 2022 doi:10.20944/preprints202206.0137.v1

these model-sharing guidelines and standards can be readily adapted for epidemiological models. We present a proof-of-concept by rebuilding 24 previously published COVID-19 transmission models using the systems biology community guidelines and infrastructures. We hope that this effort will be well received by the epidemiological modeling community and can serve as a roadmap to improve reproducibility and reusability in epidemiological modeling.

2. FAIR sharing of models

Since the SIR model was introduced in 1927, epidemiological models have grown in complexity. The programming tools used to handle these models have also diversified alongside the emergence of programming packages in different languages like R and Python, and very specific software tools such as the Spatiotemporal Epidemiological Modeler (STEM) (13) or the Epidemiological Modeling (EMOD) software (14). Earlier surveys of the field indicated that 30% of the publications did not detail the steps taken to perform the statistical analyses while 70% specified the use of a specific software package. But, in most cases, the software or packages used were not reported to be available. This has unintentionally led to the accumulation of multiple incompatibilities (10).

Mathematical models in systems biology also have become complex through the rise of interdisciplinary research. Early on the modeling community, therefore, adopted the FAIR principles for data stewardship – Findable, Accessible, Interoperable, and Reusable – to act as a guide for creation and sharing of reusable models (15). The epidemiological health researchers have more recently looked into FAIR aspects to improve the findability and accessibility of COVID-19 data (16), but not models. However, the FAIR best practices, guidelines, and implementations established in systems biology can easily be applied to the epidemiological models (Box1).

Box 1: Steps to share an epidemiological model using FAIR guidelines

- To increase findability, create a permanent and persistent digital identifier (ID) unique to the model by submitting the model to a database such as BioModels (17). This will facilitate easy identification of the model and its variants and versions.
- To facilitate accessibility, employ standardized protocols which are free and universally implementable to retrieve the model's data or metadata, as available in the BioModels repository. The protocols are also required to employ authorization mechanisms where necessary.
- To ensure interoperability, encode models in a formal, machine-readable format such as the Systems Biology Markup Language (SBML) (18), encode simulation conditions in the Simulation Experiment Description Markup Language (SED-ML) (19), and distribute the codes packed together as COMBINE archives (20).
- To promote reusability, semantically enrich models with standard meta-data that specifies their properties and attributes in an unambiguous manner; use suitable ontology terms and cross-references using identifiers.org (21), a perennial URL provider for annotation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 June 2022 doi:10.20944/preprints202206.0137.v1

3. Improving model interoperability and reusability

To comply with FAIR principles, at first, the model codes should be encoded in standardized formats to promote interoperability and then semantically enriched with relevant annotations to document model components and eliminate ambiguity in their description. Systems biology modelers have formulated and used standard languages like SBML (18) and CellML (22) to encode models, and SED-ML (19) to encode simulation conditions. These languages support a wide range of mathematical frameworks and modeling approaches including epidemiological compartmental models. These languages were also developed to handle RDF-based annotations in accordance with the Minimal Information Required to Annotate Model (MIRIAM) guidelines (23) and hence can be used to describe the context of the model.

SBML is a community standard developed under the umbrella of the Computational Modeling in Biology Network (COMBINE) (24) and is widely used by modelers to encode, publish, and exchange models. Several tools support SBML models including standalone modeling tools with user-friendly GUIs like COPASI (25), VCell (26), and Cell Designer (27). SBML can also be imported and interconverted into a human-readable text-based language called Antimony and handled using command-line tool Tellurium in a Python environment (28). SBML models can be built, manipulated, simulated, analyzed and/or visualized in many of these tools. Other popular simulation platforms like R (29), MATLAB (30), and Mathematica (31) also have dedicated libraries to handle SBML files and run simulations. With high interoperability, an SBML model can be readily imported into a wide range of supporting software and further utilized by the modeler, thereby allowing straightforward reproducibility assessment of the model. Epidemiological models of varying structural complexity can be encoded in SBML and be run with the existing infrastructure, which leads to increased interoperability and reusability. However, there are still open opportunities for further development of SBML; one example is the support for contact matrices in epidemiological models, without which multiple equations with individual contact parameters have to be encoded. Parameter estimation can be performed using SBML-supporting tools such as COPASI, MATLAB, D2D, dMOD, parPE, pyABC and pyPESTO (32). Most of the tools described in this article are also freely available through BioSimulators (https://biosimulators.org), a central registry of biosimulation tools to facilitate reproducible simulation (33). Hence, adapting SBML for epidemiological modeling will provide an added advantage in addition to tackling several of the issues which are impediments to model reproducibility and repurposing.

To enable sharing of models and virtual experiments through a single file, the Open Modelling EXchange (OMEX) (20) format was developed, and it became the basis for the COMBINE archive. A COMBINE archive is a compressed file that contains all files necessary to reproduce simulation studies. These files may be model files, simulation setups, semantic annotations, descriptions of graphical network layouts, result tables, etc. Since the development of the COMBINE archive, several software tools capable of creating and importing the OMEX file format have been deployed.

4. Improving model findability and accessibility

While encoding models in standard formats like SBML improves interoperability and reusability, it is also essential to make these models easily findable and accessible. The systems biology community addressed this requirement by setting up several public repositories for model sharing. BioModels is one of the early repositories established in 2005 and was among the first to endorse and accept submissions in the SBML format. It can host submissions from different modeling approaches in any format including the R, C++, Python, MATLAB, Mathematica, SBML and OMEX formats (34). By hosting over 2400 models from peer-reviewed literature among which about 1050 are manually curated, BioModels has increased the accessibility, reproducibility, and reusability of models in the systems biology community (35). All models submitted to BioModels generate a unique and persistent model identifier. One major reason for the enduring and expanding

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 June 2022 doi:10.20944/preprints202206.0137.v1

popularity of BioModels is its model curation service. The semantic annotation added during model curation further enables sophisticated mining of models using keyword-based search and facets-based filtering through BioModel's web interface and programmatic access. The curation pipeline at BioModels ensures the availability of reproducible and reusable model files in the public domain that can be easily accessed using the submission's unique identifier. All the models in BioModels are freely accessible under the CC0 license.

5. Manual curation of models in BioModels

At present, BioModels is one of the largest repositories of curated ODE models of biological processes. Models from direct submissions or from public literature covering a targetted area (e.g. COVID19 transmission, neurodegeneration, or diabetes) are manually curated by the BioModels team (Figure 1). Model curation involves two steps i) reproducing simulation results from the original manuscript, and ii) semantic enrichment of the models using controlled vocabularies. In the first step of the curation process, a curator examines the peer-reviewed publication associated with the model, encodes the mathematical equations in a standard format like SBML, simulates it using a tool different from the original one used by the authors, reproduces the results, and visualizes the output. Modeling tools such as COPASI, Matlab Simbiology, and Tellurium are often used to encode and simulate kinetic models. At least one of the simulation figures in the publication should be reproduced to consider a model as curated. When the simulation is reproduced, the curator semantically enriches the model by adding annotations to unambiguously indicate all the entities in the model. This involves adding machine-readable crossreferences to standard data resources, ontologies, reference publications, etc following MIRIAM guidelines (23). The curator also creates files in the OMEX and SBML formats and makes all these files publicly available and tags the submission as curated. Malik-Sheriff *et al.* 2020 (17) provide a detailed description of the model curation process with an extended list of controlled vocabularies used for annotation.

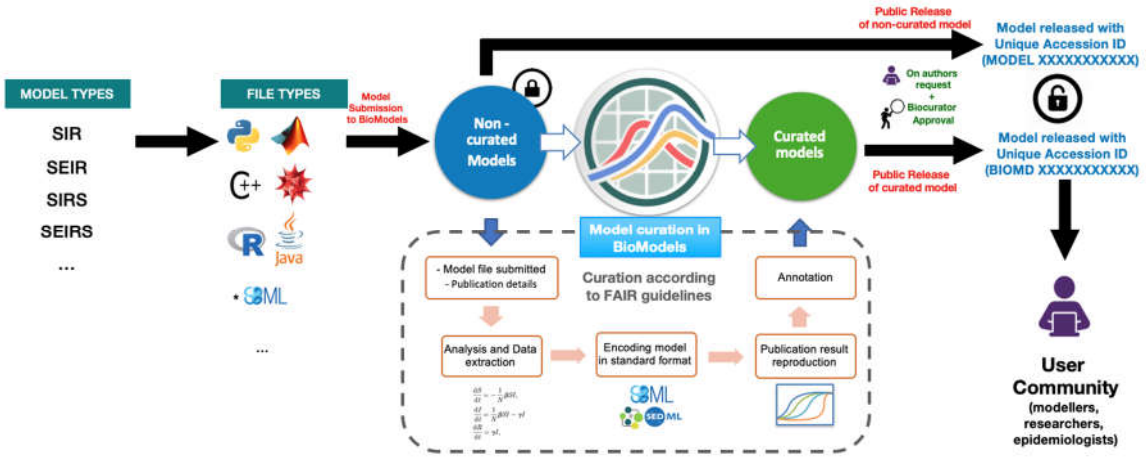

Figure 1. Model submission and FAIR guidelines-compliant curation pipeline in the BioModels: Currently BioModels accepts submissions of SIR models and its derivative versions in any file format. Models submitted can be kept private and released when the associated publication becomes publicly available. The submitted models are manually curated following FAIR guidelines before the model. If the submission does not include a standard file format, the model is re-encoded in SBML and SED-ML using tools like COPASI and Tellurium and semantically enriched with controlled vocabularies. At least one simulation figure is reproduced. All model codes and associated files are compressed into a COMBINE archive to enable single file dissemination via the BioModels platform.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 June 2022 doi:10.20944/preprints202206.0137.v1

6. FAIR COVID-19 model collection

Epidemiological models of COVID-19 capture the transmission dynamics driven by interactions between individuals of differing disease statuses. Following the FAIR principles and the curation pipeline as described above, we manually curated 24 COVID-19 transmission models from peer-reviewed, published manuscripts (Table 1). These models were carefully examined, re-encoded in the SBML format, and simulated using COPASI. If at least one of the simulation figures from the original model publication could be reproduced, the SBML model was semantically enriched with domain-specific controlled vocabularies (36). The SBML model file, its associated SED-ML simulation setups, and other supporting files were then bundled in a COMBINE archive and published openly in the BioModels with a permanent, unique identifier. Our COVID-19 collection page provides access to the full list of models (https://www.ebi.ac.uk/biomodels/search?query=submitter_keywords:COVID-19&domain=biomodels) implemented in this study The verified reproducible SBML code and the COMBINE archives of these models provide higher interoperability, thus benefiting a broad scientific community that uses diverse programming environments.

Table 1. - List of curated FAIR COVID-19 models. (a) number of species (entities) in SBML model, (b) number of global parameters (c) number of reactions in SBML model, (d) reproducibility scorecard based score for each model before curation.

Model Publication on BioModels	Model ID	No. of Compartments (a)	No. of Parameters (b)	No. of Transmission steps (c)	Reproducibility Scorecard (d)	Model codes shared publicly	Ref.
Giordano2020 - SIDARTHE model of COVID-19 spread in Italy	BIOMD0000000955	8	22	13	4	Yes	(39)
Bertozzi2020 - SIR model of scenarios of COVID-19 spread in CA and NY	BIOMD0000000956	3	21	2	4	Yes	(40)
Roda2020 - SIR model of COVID-19 spread in Wuhan	BIOMD0000000957	4	3	3	3	No	(41)
Ndairou2020 - early-stage transmission dynamics of COVID-19 in Wuhan	BIOMD0000000958	8	14	12	3	No	(42)
Paiva2020 - SEIAHRD model of transmission dynamics of COVID-19	BIOMD0000000960	8	133	11	2	No	(43)
Zhao2020 - SUQC model of COVID-19 transmission dynamics in Wuhan, Hubei, and China	BIOMD0000000962	5	37	3	3	No	(44)
Weitz2020 - SIR model of COVID-19 transmission with shielding	BIOMD0000000963	3	5	2	4	Yes	(45)
Mwalili2020 - SEIR model of COVID-19 transmission and environmental pathogen prevalence	BIOMD0000000964	6	15	14	3	No	(46)
Cuadros2020 - SIHRD spatiotemporal model of COVID-19 transmission in Ohio	BIOMD0000000969	28	36	32	3	No	(47)
Hou2020 - SEIR model of COVID-19 transmission in Wuhan	BIOMD0000000970	5	6	3	3	No	(48)
Tang2020 - Estimation of transmission risk of COVID-19 and impact of public health interventions	BIOMD0000000971	8	16	13	3	No	(49)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 June 2022 doi:10.20944/preprints202206.0137.v1

Tang2020 - Estimation of transmission risk of COVID-19 and impact of public health interventions - updated	BIOMD0000000972	8	22	13	3	No	(49)
Carcione2020 - Deterministic SEIR simulation of a COVID-19 outbreak	BIOMD0000000974	6	6	4	4	Yes	(50)
Ghanbari2020 - forecasting the second wave of COVID-19 in Iran	BIOMD0000000976	4	7	5	3	No	(51)
Sarkar2020 - SAIR model of COVID-19 transmission with quarantine measures in India	BIOMD0000000977	6	13	15	3	No	(52)
Mukandavire2020 - SEIR model of early COVID-19 transmission in South Africa	BIOMD0000000978	4	7	3	3	No	(53)
Malkov2020 - SEIRS model of COVID-19 transmission with reinfection	BIOMD0000000979	5	6	4	3	No	(54)
Malkov2020 - SEIRS model of COVID-19 transmission with time-varying R values and reinfection	BIOMD0000000980	5	13	4	3	No	(54)
Wan2020 - risk estimation and prediction of the transmission of COVID-19 in maninland China excluding Hubei province	BIOMD0000000981	12	17	15	3	No	(55)
Law2020 - SIR model of COVID-19 transmission in Malyasia with time-varying parameters	BIOMD0000000982	3	10	2	3	No	(56)
Zongo2020 - model of COVID-19 transmission dynamics under containment measures in France	BIOMD0000000983	7	21	10	3	No	(57)
Fang2020 - SEIR model of COVID-19 transmission considering government interventions in Wuhan	BIOMD0000000984	5	5	3	3	No	(58)
Westerhoff2020 - systems biology model of the coronavirus pandemic 2020	BIOMD0000000988	20	26	18	3	No	(59)
Okuonghae2020 - SEAIR model of COVID-19 transmission in Lagos, Nigeria	BIOMD0000000991	9	17	10	3	No	(60)

The epidemiological forecasting models in our collection provided insights into how disease transmission occurs in the real world and helped to assess the impact of non-pharmaceutical interventions to control disease spread. The simplest model of this kind is the SIR model with three compartments. For the COVID-19 pandemic, several models have been proposed based on the SIR model with varying degrees of complexity to account for the asymptomatic spreaders, differing social mixing patterns, varying demographics, and government responses. From the models we encoded, three models were simple SIR models, one model was a SIR model with time-dependent transmission parameters. Seven were SEIR models with an additional "Exposed (E)" compartment. The "Exposed" compartment was used in the model to account for the time delay in detecting new infections. The remaining models were further modifications of the SIR models taking into account the effect of asymptomatic spread, contact tracing, forecasting hospital capacities, understanding the effect of geographical proximity to transport nodes like airports on transmission, and the scenario of reinfection for a second wave. With further submissions, we hope to rapidly expand our collection of reproducible and reusable COVID-19 models which are already utilized by the community.

 doi:10.20944/preprints202206.0137.v1

7. Reproducibility scorecard for models

A cornerstone of scientific research is the ability to reproduce results. Reproducibility is the ability to construct the code *de novo* and/or ensure that the mathematical expressions are properly represented and reproduce the simulation results in software different from the one originally used (37). Ensuring model reproducibility will help increase the reliability and ease of reusability of the model. In our recent study (8) we discussed the implications of the observed reproducibility crisis and proposed an 8-point reproducibility scorecard (38) for modelers, reviewers, and journal editors to evaluate models during the peer-review process. We showed that an increase in score increased the chance of model reproducibility and recommended a minimal cut-off of 4. A majority of the COVID-19 models we curated received a score of 3 out of 8 before curation and took significant effort to re-encode and reproduce them. It is therefore clear that a straightforward and explicit recording of the assumptions as outlined by the EPIFORGE guidelines when used alongside the reproducibility scorecard would ensure the mathematical reproducibility and contextual reliability of epidemiological models. Such a comprehensive, combined checklist will standardize reporting formats, guarantee model reproducibility, increase model findability, and ensure ease of model repurposing and reuse.

8. Conclusion

COVID-19 is the first pandemic of this scale in over a century, and its progression was closely studied, monitored, and modeled across the globe. We hope that our effort to implement the standards developed by the systems biology community in COVID-19 epidemiological models can serve as a roadmap to enhance the findability, accessibility, interoperability, and reusability of the models. Following the FAIR guidelines will increase the reproducibility and repurposing of these epidemiological models and help us stay prepared for pandemics. Additionally, models in the SBML, or similar domain-specific formats can benefit from tools and algorithms used for model fitting, parameter estimation, sensitivity, and identifiability analysis(18). Epidemiological modelers can benefit from reusing existing modeling packages to build and analyze models. SBML format specifications and supporting libraries including libSBML and JSBML are openly available and can be used to implement SBML-support in epidemiological software for modeling and simulation.

A recent funding call from US Defense Advanced Research Projects Agency (DARPA) titled "Automating Scientific Knowledge Extraction and Modeling (ASKEM)" invited proposals to develop approaches and tools for machine-assisted modeling and simulation through a formal representation of epidemiological models of viral pandemics like COVID-19 (1). The representation of COVID-19 models in the machine-readable SBML format would profoundly facilitate such development. By adopting a common format, systems biology and epidemiological modeling communities can greatly benefit from each other's modeling infrastructures. Although our study discussed model FAIRifcation in the context of COVID-19 models, our recommendations are applicable to all epidemiological disease transmission models. In conjunction with the EPIFORGE guidelines, we recommend the use of the reproducibility scorecard to assess models' FAIRness during the peer-review process and thereby improve the reproducibility, reliability, and reusability of the COVID-19 models.

Acknowledgments: MK was supported by the Federal Ministry of Education and Research (BMBF, Germany) within the research network Systems Medicine of the Liver (LiSyM, grant number 031L0054), by the German Research Foundation (DFG) within the Research Unit Programme FOR 5151 "QuaLiPerF (Quantifying Liver Perfusion-Function Relationship in Complex Resection - A Systems Medicine Approach)" by grant number 436883643. KR, HH, and RSMS are supported by EMBL core funding and would like to acknowledge the funding of EMBL member states. The authors received funding for the creation of the COVID-19 model collection from the EOSC secretariat.eu which received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-05-2018-2019, Grant Agreement number 831644.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. DARPA 2021. DARPA 2021. DARPA: Automating Scientific Knowledge Extraction and Modeling (ASKEM), Retrieved from the web 9 Mar, 2022. https://sam.gov/api/prod/opps/v3/opportunities/resources/files/c4885d3cea7c40cd96ec8db407f4934f/download?
2. Pollett S, Johansson MA, Reich NG, Brett-Major D, Del Valle SY, Venkatramanan S, et al. Recommended reporting items for epidemic forecasting and prediction research: The EPIFORGE 2020 guidelines. PLoS Med. 2021 Oct;18(10):e1003793.
3. Kermack WO, McKendrick AG, Walker GT. A contribution to the mathematical theory of epidemics. Proc R Soc Lond Ser Contain Pap Math Phys Character. 1927 Aug 1;115(772):700–21.
4. Adam D. Special report: The simulations driving the world's response to COVID-19. Nature. 2020;580(7802):316–9.
5. Jalali MS, DiGennaro C, Sridhar D. Transparency assessment of COVID-19 models. Lancet Glob Health. 2020 Dec 1;8(12):e1459–60.
6. Rogers LS, Health JBS of P. Transparency, Reproducibility, and Validation of COVID-19 Projection Models [Internet]. Johns Hopkins Bloomberg School of Public Health. [cited 2021 Jul 15]. Available from: https://www.jhsph.edu/covid-19/articles/transparency-reproducibility-and-validation-of-covid-19-projection-models.html
7. Becker AD, Grantz KH, Hegde ST, Bérubé S, Cummings DAT, Wesolowski A. Development and dissemination of infectious disease dynamic transmission models during the COVID-19 pandemic: what can we learn from other pathogens and how can we move forward? Lancet Digit Health. 2021 Jan 1;3(1):e41–50.
8. Tiwari K, Kananathan S, Roberts MG, Meyer JP, Sharif Shohan MU, Xavier A, et al. Reproducibility in systems biology modelling. Mol Syst Biol [Internet]. 2021 Feb [cited 2022 Jan 10];17(2). Available from: https://onlinelibrary.wiley.com/doi/10.15252/msb.20209982
9. Kirouac DC, Cicali B, Schmidt S. Reproducibility of Quantitative Systems Pharmacology Models: Current Challenges and Future Opportunities. CPT Pharmacomet Syst Pharmacol. 2019 Apr;8(4):205–10.
10. Peng RD, Dominici F, Zeger SL. Reproducible Epidemiologic Research. Am J Epidemiol. 2006 May 1;163(9):783–9.
11. Bierer BE, Li R, Barnes M, Sim I. A Global, Neutral Platform for Sharing Trial Data. N Engl J Med. 2016 Jun 23;374(25):2411–3.
12. Shepherd BE, Blevins Peratikos M, Rebeiro PF, Duda SN, McGowan CC. A Pragmatic Approach for Reproducible Research With Sensitive Data. Am J Epidemiol. 2017 Aug 15;186(4):387–92.
13. Douglas JV, Bianco S, Edlund S, Engelhardt T, Filter M, Günther T, et al. STEM: An Open Source Tool for Disease Modeling. Health Secur. 2019 Aug 1;17(4):291–306.
14. Bershteyn A, Gerardin J, Bridenbecker D, Lorton CW, Bloedow J, Baker RS, et al. Implementation and applications of EMOD, an individual-based multi-disease modeling platform. Pathog Dis. 2018 Jul 6;76(5):fty059.
15. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3(1):160018.
16. NFDI4Health Task Force COVID-19, Darms J, Henke J, Hu X, Schmidt CO, Golebiewski M, et al. Improving the FAIRness of Health Studies in Germany: The German Central Health Study Hub COVID-19. In: Delgado J, Benis A, de Toledo P, Gallos P, Giacomini M, Martínez-García A, et al., editors. Studies in Health Technology and Informatics [Internet]. IOS Press; 2021 [cited 2022 Jun 2]. Available from: https://ebooks.iospress.nl/doi/10.3233/SHTI210818
17. Malik-Sheriff RS, Glont M, Nguyen TVN, Tiwari K, Roberts MG, Xavier A, et al. BioModels—15 years of sharing computational models in life science. Nucleic Acids Res. 2019;48(D1):D407–15.
18. Keating SM, Waltemath D, König M, Zhang F, Dräger A, Chaouiya C, et al. SBML Level 3: an extensible format for the exchange and reuse of biological models. Mol Syst Biol. 2020 Aug;16(8):e9110.
19. Waltemath D, Adams R, Bergmann FT, Hucka M, Kolpakov F, Miller AK, et al. Reproducible computational biology experiments with SED-ML - The Simulation Experiment Description Markup Language. BMC Syst Biol. 2011 Dec;5(1):198.
20. Bergmann FT, Adams R, Moodie S, Cooper J, Glont M, Golebiewski M, et al. COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. BMC Bioinformatics. 2014 Dec 14;15(1):369.
21. Juty N, Le Novère N, Laibe C. Identifiers.org and MIRIAM Registry: community resources to provide persistent identification. Nucleic Acids Res. 2012 Jan;40(Database issue):D580-586.
22. Lloyd CM, Halstead MDB, Nielsen PF. CellML: its future, present and past. Prog Biophys Mol Biol. 2004 Jul;85(2–3):433–50.
23. Le Novère N, Finney A, Hucka M, Bhalla US, Campagne F, Collado-Vides J, et al. Minimum information requested in the annotation of biochemical models (MIRIAM). Nat Biotechnol. 2005 Dec;23(12):1509–15.
24. Waltemath D, Golebiewski M, Blinov ML, Gleeson P, Hermjakob H, Hucka M, et al. The first 10 years of the international coordination network for standards in systems and synthetic biology (COMBINE). J Integr Bioinforma. 2020 Jun 29;17(2–3).
25. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI—a COmplex PAthway SImulator. Bioinformatics. 2006 Dec 15;22(24):3067–74.
26. Moraru II, Schaff JC, Slepchenko BM, Loew LM. The virtual cell: an integrated modeling environment for experimental and computational cell biology. Ann N Y Acad Sci. 2002 Oct;971:595–6.
27. Funahashi A, Morohashi M, Kitano H, Tanimura N. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. BIOSILICO. 2003 Nov;1(5):159–62.

28. Choi K, Medley JK, König M, Stocking K, Smith L, Gu S, et al. Tellurium: An Extensible Python-based Modeling Environment for Systems and Synthetic Biology. Biosystems. 2018 Sep;171:74–9.
29. Radivoyevitch T. A two-way interface between limited Systems Biology Markup Language and R. BMC Bioinformatics. 2004 Dec 7;5(1):190.
30. Keating SM, Bornstein BJ, Finney A, Hucka M. SBMLToolbox: an SBML toolbox for MATLAB users. Bioinformatics. 2006 May 15;22(10):1275–7.
31. Shapiro BE, Hucka M, Finney A, Doyle J. MathSBML: a package for manipulating SBML-based biological models. Bioinforma Oxf Engl. 2004 Nov 1;20(16):2829–31.
32. Schmiester L, Schälte Y, Bergmann FT, Camba T, Dudkin E, Egert J, et al. PEtab—Interoperable specification of parameter estimation problems in systems biology. Schneidman-Duhovny D, editor. PLOS Comput Biol. 2021 Jan 26;17(1):e1008646.
33. Shaikh B, Smith LP, Vasilescu D, Marupilla G, Wilson M, Agmon E, et al. BioSimulators: a central registry of simulation engines and services for recommending specific tools. Nucleic Acids Res. 2022 May 7;gkac331.
34. Glont M, Nguyen TVN, Graesslin M, Hälke R, Ali R, Schramm J, et al. BioModels: expanding horizons to include more modelling approaches and formats. Nucleic Acids Res. 2018 Jan 4;46(D1):D1248–53.
35. Mendes P. Reproducible Research Using Biomodels. Bull Math Biol. 2018 Dec;80(12):3081–7.
36. Courtot M, Juty N, Knüpfer C, Waltemath D, Zhukova A, Dräger A, et al. Controlled vocabularies and semantics in systems biology. Mol Syst Biol. 2011 Oct 25;7:543.
37. Miłkowski M, Hensel WM, Hohol M. Replicability or reproducibility? On the replication crisis in computational neuroscience and sharing only relevant detail. J Comput Neurosci. 2018 Dec 1;45(3):163–72.
38. Tiwari K, Waltemath D, Hermjakob H, Malik-Sheriff RS. Reproducibility Scorecard to Assess Systems Biology Models [Internet]. Zenodo; 2021. Available from: https://doi.org/10.5281/zenodo.5786693
39. Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nat Med. 2020 Jun;26(6):855–60.
40. Bertozzi AL, Franco E, Mohler G, Short MB, Sledge D. The challenges of modeling and forecasting the spread of COVID-19. Proc Natl Acad Sci U S A. 2020 Jul 21;117(29):16732–8.
41. Roda WC, Varughese MB, Han D, Li MY. Why is it difficult to accurately predict the COVID-19 epidemic? Infect Dis Model. 2020;5:271–81.
42. Ndaïrou F, Area I, Nieto JJ, Torres DFM. Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. Chaos Solitons Fractals. 2020 Jun;135:109846.
43. Paiva HM, Afonso RJM, de Oliveira IL, Garcia GF. A data-driven model to describe and forecast the dynamics of COVID-19 transmission. PloS One. 2020;15(7):e0236386.
44. Zhao S, Chen H. Modeling the epidemic dynamics and control of COVID-19 outbreak in China. Quant Biol Beijing China. 2020 Mar 11;1–9.
45. Weitz JS, Beckett SJ, Coenen AR, Demory D, Dominguez-Mirazo M, Dushoff J, et al. Modeling shield immunity to reduce COVID-19 epidemic spread. Nat Med. 2020 Jun;26(6):849–54.
46. Mwalili S, Kimathi M, Ojiambo V, Gathungu D, Mbogo R. SEIR model for COVID-19 dynamics incorporating the environment and social distancing. BMC Res Notes. 2020 Jul 23;13(1):352.
47. Cuadros DF, Xiao Y, Mukandavire Z, Correa-Agudelo E, Hernández A, Kim H, et al. Spatiotemporal transmission dynamics of the COVID-19 pandemic and its impact on critical healthcare capacity. Health Place. 2020 Jul;64:102404.
48. Hou C, Chen J, Zhou Y, Hua L, Yuan J, He S, et al. The effectiveness of quarantine of Wuhan city against the Corona Virus Disease 2019 (COVID-19): A well-mixed SEIR model analysis. J Med Virol. 2020 Jul;92(7):841–8.
49. Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. J Clin Med. 2020 Feb 7;9(2):E462.
50. Carcione JM, Santos JE, Bagaini C, Ba J. A Simulation of a COVID-19 Epidemic Based on a Deterministic SEIR Model. Front Public Health. 2020;8:230.
51. Ghanbari B. On forecasting the spread of the COVID-19 in Iran: The second wave. Chaos Solitons Fractals. 2020 Nov;140:110176.
52. Sarkar K, Khajanchi S, Nieto JJ. Modeling and forecasting the COVID-19 pandemic in India. Chaos Solitons Fractals. 2020 Oct;139:110049.
53. Mukandavire Z, Nyabadza F, Malunguza NJ, Cuadros DF, Shiri T, Musuka G. Quantifying early COVID-19 outbreak transmission in South Africa and exploring vaccine efficacy scenarios. PloS One. 2020;15(7):e0236003.
54. Malkov E. Simulation of coronavirus disease 2019 (COVID-19) scenarios with possibility of reinfection. Chaos Solitons Fractals. 2020 Oct;139:110296.
55. Wan H, Cui JA, Yang GJ. Risk estimation and prediction of the transmission of coronavirus disease-2019 (COVID-19) in the mainland of China excluding Hubei province. Infect Dis Poverty. 2020 Aug 24;9(1):116.
56. Law KB, Peariasamy KM, Gill BS, Singh S, Sundram BM, Rajendran K, et al. Tracking the early depleting transmission dynamics of COVID-19 with a time-varying SIR model. Sci Rep. 2020 Dec 10;10(1):21721.
57. Zongo P, Zorom M, Mophou G, Dorville R, Beaumont C. A model of COVID-19 transmission to understand the effectiveness of the containment measures: application to data from France. Epidemiol Infect. 2020 Sep 22;148:e221.
58. Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. J Med Virol. 2020 Jun;92(6):645–59.

 doi:10.20944/preprints202206.0137.v1

59. Westerhoff HV, Kolodkin AN. Advice from a systems-biology model of the corona epidemics. NPJ Syst Biol Appl. 2020 Jun 12;6(1):18.
60. Okuonghae D, Omame A. Analysis of a mathematical model for COVID-19 population dynamics in Lagos, Nigeria. Chaos Solitons Fractals. 2020 Oct;139:110032.