

Standards-Based Reproducible Biochemical Modeling Checklist

Data aggregation:

- New data collection and experimental procedures are reproducible
 - The experimental protocol is provided
 - Conditions are provided
 - Measurement uncertainty is quantified
 - A record of when the data was collected is provided
 - A record of the individual or lab that produced the data is provided
- Data collected from databases and literature are curated and metadata are provided
 - Data and metadata are version-controlled using a reusable platform (e.g. Quilt, PROV)
 - Provenance information is provided
 - Provenance information follows the SEPIO
 - Data collection and metadata storage are automated

Model description:

- All species and parameter names, types of rate laws, and the roles of species in reactions use standardized ontologies (e.g. InChI, SBO, BpForms, BcForms)
- All model component descriptions and annotations meet the MIRIAM standard
- The model is developed in or exported to a machine-readable standardized modeling format (e.g. CellML, SBML)
 - A standard-compatible tool is used to write SBML or CellML (e.g. Antimony, COPASI)

Parameter estimation:

- Reusable programs are used to estimate parameters on an SBML model (e.g. COPASI, SBML-PET, PyBioNetFit)
- Confidence intervals are reported on estimated parameters and the routine used to generate confidence intervals (e.g. Monte Carlo) is reported
- Families of parameter values are reported for non-identifiable models, if there are a discrete set of possibilities
- Uncertainty quantification is performed, and the estimation routine used is provided

Simulation:

- Simulation experiments are described in SED-ML or an alternative standardized format (e.g. SESSL)
- Simulation experiments meet the MIASE guidelines
- SED-ML documents are annotated using the KiSAO
- Reusable, standard-compatible simulators are used to perform simulation experiments (e.g. COPASI, JWS Online, Tellurium/libRoadRunner)

Results storage:

- All data is provided in annotated, computer-readable spreadsheets (e.g. using ObjTables or the SEEK platform with RightField integration)
 - Unprocessed results and data included in published figures and tables are provided
 - All data fields are richly annotated using RightField
- Source data and code to produce published figures is provided, or can be automatically reproduced by uploading models and simulation experiment descriptions to JWS Online for real-time simulation and visualization

Verification & validation:

- All modeling code is systematically verified
 - Continuous integration is used to ensure changes to the model do not alter expected behavior (e.g. using platforms like CircleCI, Travis)
- The model and associated simulation experiments are validated to assess biological relevance
 - Tests are written in a reusable program using SciUnit to automatically check that all model components exhibit expected biological behavior
- The model and associated simulation experiments are tested in an independent computing environment to ensure the software is portable and reproducible (e.g. a VM, Docker image, or cloud services are used)

Documentation:

- Design decisions and assumptions are recorded as metadata used to annotate the model
- Model and associated software, code, and data are version-controlled (e.g. using Git or SVN)
- Standardized graphical notation is used to visualize the model components and their interactions for comprehension (e.g. using SBGN)

Packaging:

- Model SBML and SED-ML files are packaged in a COMBINE archive
- Complete modeling studies are packaged into a VM or Docker image that contains the operating system, dependencies, and simulators used to enable exact replication of the original modeling environment

Dissemination:

- The modeling study is disseminated in a pre-print server (e.g. bioRxiv)
- The modeling study is published in a peer-reviewed journal, preferably with open-access
- The model is available at a public repository
 - The model is uploaded in SBML format to the BioModels Database and a workspace for the model is created on JWS Online to enable real-time web-based simulations
 - The repository materials are governed by an open-source license