Leveraging AI for improved reproducibility of mathematical disease models: insights from a retinitis pigmentosa case study

Elea Thibault Greugny[1*], Nicolas Ratto[1*], Loïc Etheve[1], Jean-Baptise Gourlet[1], **Simon Arsène**[1], Jim Bosley[1], Frederic Cogny[1], Claudio Monteiro[1]

[1] Nova In Silico, Lyon, France

[*] These authors contributed equally

BACKGROUND

Mathematical modeling of disease and drug action is becoming an indispensable component of drug development, highlighted by recent examples of models predicting trial results. To be able to rely on such approaches, decision-makers need to be able to verify those results independently with in silico confirmatory studies. **Artificial Intelligence** (AI) offers a valuable avenue for **improving the reproducibility of complex mathematical models**. Here, using a case study of a disease progression mechanistic model of **retinitis pigmentosa** (RP) [1], we showcase how AI could be leveraged to **accelerate, secure and simplify the reuse process** of mathematical models.

Leveraging **large language models (LLM)** can help to **automate mathematical disease models reuse** and **improve results reproducibility**

Issues encountered for mathematical model results reproducibility

Omissions and/or errors in publication

- Initial conditions
- Parameter values
- Virtual population details
- Units
- *In silico* protocol (dose, treatment schedule)
- Missing equations terms

Errors introduced during the reuse process

- Typographical errors in parameter names or values
- Sign issues

Lengthy and cumbersome process

- If model code not provided, manual reimplementation of the model from publication.
- Translation of the code into the desired programming language.

Almost half of the published mathematical models are not directly reproducible [2]

AI as a tool to streamline the reproducibility process

Detection of missing values such as initial conditions or parameter values **during the translation process**

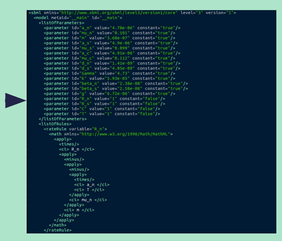
Automated translation of equations into code

$$
\begin{aligned}
R_n' &= R_n(a_n T - \mu_n - m) \\
R_s' &= R_s(a_s T - \mu_s) + m R_n \\
C' &= C(a_c T - \mu_c + d_n R_n + d_s R_s) \\
T' &= T(\Gamma - kT - \beta_n R_n - \beta_s R_s - \gamma C),
\end{aligned}
$$

Fig 1. Retinitis pigmentosa model equations from [1], describing the evolution of the number of normally functioning rods (Rn), non-functioning rods (Rs), cones (C), and the nutrient pool (T).

Chat GPT 4

MathPix



SBML Systems Biology Markup Language

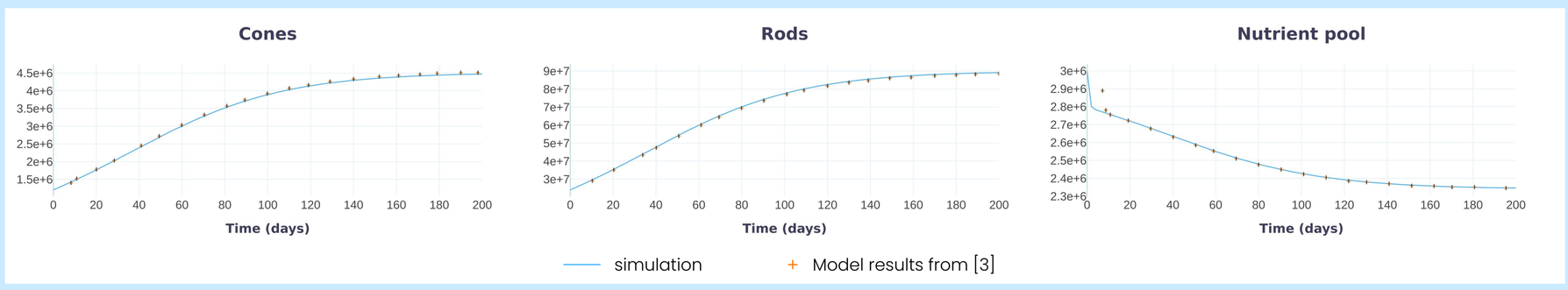
Acceleration and simplification of the process

- From publication PDF to reusable code in a few minutes

Streamlined process and early error detection allows more time for review and correction

Retinitis pigmentosa model case study

Fig 2. Recreating simulation results originally presented in [3].
The simulation results and parameter values were derived from [3]. Initial conditions had to be inferred from the graphs for simulation, leading to slight variations.



METHODS

We used Chat GPT 4, a sophisticated large language model developed by OpenAI [4], customized with Mathpix additional functionalities [5]. This customized model enabled the extraction of equations from PDF documents and their conversion into Antimony, a human-readable, text-based language.

Subsequently, the model was translated into Systems Biology Markup Language (SBML) using the Python package tellurium. The converted model was then uploaded onto the jinkō platform for simulations.

RESULTS

The efficient and accurate implementation of the RP model using AI techniques was demonstrated in this study. The disease progression biological phenomena such as the evolutions of rods, cones and the nutrient pool (representative of the total number of retinal-pigment epithelium neuroprotective factors, growth factors and nutrients) as reported in the literature was successfully reproduced.

DISCUSSION & CONCLUSION

- We demonstrate here that **AI can reduce the burden of reimplementing published models** by accelerating the process and identifying missing information. We developed a method enabling researchers to **quickly obtain reusable code in their preferred programming language while highlighting areas needing review** and allowing them to focus on complex corrections, requiring deep understanding of the model.
- While AI can optimize the process of reimplementing models, **it does not eliminate the need for rigorous review**, as results generated by AI must always be independently validated by domain experts.
- This quick and efficient method can be a **valuable tool for authors and reviewers during the publication process to ensure model reproducibility**, as well as for readers aiming to reuse the model.
- The ideal solution for ensuring model reproducibility remains **direct access to the original model code** used to generate the published results.

REFERENCES

[1] Camacho, E. T., Punzo, C. & Wirkus, S. A. Quantifying the metabolic contribution to photoreceptor death in retinitis pigmentosa via a mathematical model. Journal of Theoretical Biology vol. 408 75–87 (2016).

[2] Tiwari K, Kananathan S et al. Reproducibility in systems biology modelling. Molecular Systems Biology vol. 17,2 (2021).

[3] Camacho ET, Wirkus S. Tracing the progression of retinitis pigmentosa via photoreceptor interactions. Journal of Theoretical Biology vol. 317 105-118 (2013).

[4] OpenAI. (2024). ChatGPT (April 2024 version) [Large language model]. https://chat.openai.com/chat

[5] Xiufeng Liu, Mathpix GPT. (2024) [Online] https://chatgpt.com/g/g-lwSvMnJvI-mathpix