

Nicolas Ratto¹, Jean-Baptiste Gourlet¹, Matthieu Coudron¹, Jean Ponchon¹, Nicolas Duteuil¹, Jérémy Villard¹, Loïc Etheve¹,

Claudio Monteiro¹

¹ Nova In Silico, Lyon, France

BACKGROUND

Pharmacometrics (PMx) models are often stored in unstructured PDFs or proprietary source files, which makes them difficult to reuse, maintain, or integrate with Quantitative Systems Pharmacology (QSP) models. This hampers reproducibility and slows innovation. To address this challenge, we have integrated AI-powered tools into the Jinkō platform to automate the extraction and standardization of PMx models, thereby transforming legacy content into a digital, interoperable and reusable library.

Jinkō QSP platform allows for streamlined integration of literature models



Identifying ODEs within a paper

Second-generation minimal physiologically-based pharmacokinetic model for monoclonal antibodies

Yanguang Cao¹, Joseph P Balthasar¹, and William J Jusko^{1,2}

¹Department of Pharmaceutical Sciences, School of Pharmacy and Pharmaceutical Sciences, State University of New York at Buffalo, Buffalo, NY, 14214, USA

Abstract

About half of native IgG was found distributed in extravascular space [12] and interstitial fluid (ISF) is generally believed to be the primary extravascular distributable space given the limited cellular permeability of mAb. The vascular endothelial endosome is frequently included in full-PBPK models to assess the influence of FcRn salvage [10, 15]. However, it is known that an adult has about 1.0 kg of endothelial cells, which comprise around 1.6×10^{12} cells with a surface area of 1.7 m^2 [16]. Given that the endosome is only a small part of cell space (about 6%–25%) [17, 18], the overall volume of vascular endothelial endosomes in adults may be estimated to be only about 0.4% of plasma and less than 0.1% of ISF volume. As such, even if endosomes have slightly higher concentrations of mAb than plasma, only a small fraction of the total mass of antibody will reside within vascular endothelial endosomes at any moment of time. The interaction between FcRn and antibody within the endosomal space clearly plays an important role in mAb elimination (and IgG homeostasis); however, this may be explained by the rapid time course of endosomal transit (described in detail in the full PBPK model of Chen and Balthasar [19]). Given that extravascular distribution appears to be primarily determined by mAb in ISF, with little contribution relating to the mass of mAb within endothelial endosomes, the mPBPK model takes only ISF into consideration as the extravascular distribution space.

The model structure is shown in Figure 1. Model A is described by the differential equation structure similar to full-PBPK models:

$$\frac{dC_p}{dt} = \frac{\text{Input}}{V_p} - [C_{\text{lymph}} \bullet L - C_p \bullet L_1 \bullet (1 - \sigma_1) - C_p \bullet L_2 \bullet (1 - \sigma_2) - C_{\text{leaky}}]/V_p \quad (1)$$

$$\frac{dC_{\text{tight}}}{dt} = [L_1 \bullet (1 - \sigma_1) \bullet C_p - L_1 \bullet (1 - \sigma_1) \bullet C_{\text{tight}}]/V_{\text{tight}} \quad (2)$$

$$\frac{dC_{\text{leaky}}}{dt} = [L_2 \bullet (1 - \sigma_2) \bullet C_p - L_2 \bullet (1 - \sigma_2) \bullet C_{\text{leaky}}]/V_{\text{leaky}} \quad (3)$$

$$\frac{dC_{\text{lymph}}}{dt} = [L_1 \bullet (1 - \sigma_1) \bullet C_{\text{tight}} + L_2 \bullet (1 - \sigma_2) \bullet C_{\text{leaky}} - C_{\text{lymph}} \bullet L]/V_{\text{lymph}} \quad (4)$$

where C_p is mAb concentration in V_p (plasma volume), C_{tight} and C_{leaky} are mAb ISF concentrations in tissue V_{tight} and V_{leaky} . The V_{tight} and V_{leaky} are volumes of ISF in tissues that have continuous and discontinuous or fenestrated capillaries. Based on the report of Sarin [20], we have assigned the muscle, skin, adipose and brain to V_{tight} and all other tissues to V_{leaky} (liver, kidney, heart, etc). The V_{lymph} is lymph volume, assumed equal to blood volume [21]. The L is total lymph flow equal to the sum of L_1 and L_2 , the lymph flow for V_{tight} and V_{leaky} . The σ_1 and σ_2 are vascular reflection coefficients for V_{tight} and V_{leaky} .

“Second-generation minimal physiologically-based pharmacokinetic model for monoclonal antibodies”

Cao Y, Balthasar JP, Jusko WJ. *J Pharmacokinet Pharmacodyn.* 2013 Oct;40(5):597–607.

Extraction Workflow:

- Optical Recognition (OCR):** Converted PDF content into machine-readable text, capturing mathematical expressions.
- Large Language Model (LLM) Processing:** Refined OCR outputs to reconstruct equations in LaTeX format, ensuring accuracy and semantic clarity.

Creation of the Computational Model

Step 1: ODEs extraction

- Conversion of the LaTeX equations into jinkō compatible syntax
- A description is automatically added

$$\frac{dC_p}{dt} = \frac{\text{Input}}{V_p} + [C_{\text{lymph}} \bullet L - C_p \bullet L_1 \bullet (1 - \sigma_1) - C_p \bullet L_2 \bullet (1 - \sigma_2) - C_{\text{leaky}}]/V_p \quad (1)$$

$$\frac{dC_{\text{tight}}}{dt} = [L_1 \bullet (1 - \sigma_1) \bullet C_p - L_1 \bullet (1 - \sigma_1) \bullet C_{\text{tight}}]/V_{\text{tight}} \quad (2)$$

$$\frac{dC_{\text{leaky}}}{dt} = [L_2 \bullet (1 - \sigma_2) \bullet C_p - L_2 \bullet (1 - \sigma_2) \bullet C_{\text{leaky}}]/V_{\text{leaky}} \quad (3)$$

$$\frac{dC_{\text{lymph}}}{dt} = [L_1 \bullet (1 - \sigma_1) \bullet C_{\text{tight}} + L_2 \bullet (1 - \sigma_2) \bullet C_{\text{leaky}} - C_{\text{lymph}} \bullet L]/V_{\text{lymph}} \quad (4)$$

Step 2: Parameters extraction

- Parameters used in the ODEs are recursively searched in both the text and tables
- Description, value and units are automatically added

Volume, L is total lymph volume, and:

Compound	Model A: CL from plasma				Model B: CL from ISF			
	Q_1 (%)	Q_2 (%)	CL_p (L/hr)	α_{ph}	Q_1 (%)	Q_2 (%)	CL_t (L/hr)	α_{ph}
Adcanumab	0.883 (8.64)	0.524 (7.37)	0.0000 (1.19)	6266	0.416 (44.8)	0.296 (33.1)	0.0318 (18.29)	6238
Mepolizumab	0.950 (-) ^c	0.750 (1.48)	0.0051 (1.50)	1327	0.950 (-)	0.591 (1.96)	0.0193 (2.44)	938
Gevokizumab	0.931 (2.58)	0.837 (2.63)	0.0068 (1.87)	309	0.757 (6.82)	0.834 (4.87)	0.0193 (4.53)	322
GNAC1	0.915 (5.21)	0.831 (5.71)	0.00714 (2.93)	1574	0.776 (15.2)	0.787 (11.7)	0.0189 (6.72)	1381
MEB-528	0.987 (6.17)	0.754 (7.90)	0.0537 (6.86)	2284	0.942 (10.9)	0.735 (18.3)	0.0162 (9.62)	2214
Tefibazumab	0.902 (7.37)	0.815 (7.56)	0.0033 (2.39)	6805	0.844 (23.8)	0.819 (14.0)	0.0215 (8.25)	5949
PAnAb	0.959 (-)	0.779 (5.24)	0.0067 (2.29)	68643	0.950 (-)	0.621 (3.16)	0.0224 (5.29)	52639
PR05780	0.984 (22.2)	0.638 (19.4)	0.0124 (18.3)	25967	0.575 (48.4)	0.702 (25.5)	0.0189 (32.7)	27601
Silixumab	0.964 (10.3)	0.673 (9.27)	0.0115 (5.19)	184681	0.950 (-)	0.467 (6.41)	0.0218 (18.25)	179121
Silixumab	0.949 (4.63)	0.834 (5.44)	0.0152 (3.02)	4419	0.567 (34.0)	0.806 (17.1)	0.0380 (6.93)	4433

^a for mAb and other parameters, change will likely be 0.4% (6) and (7), labeled based on L_1 (0.67) and L_2 (0.67) are listed in med as 2.9 L.

Step 3: Traceability of the model

- Links are added to ODEs and parameters components to track them back to the source in the paper

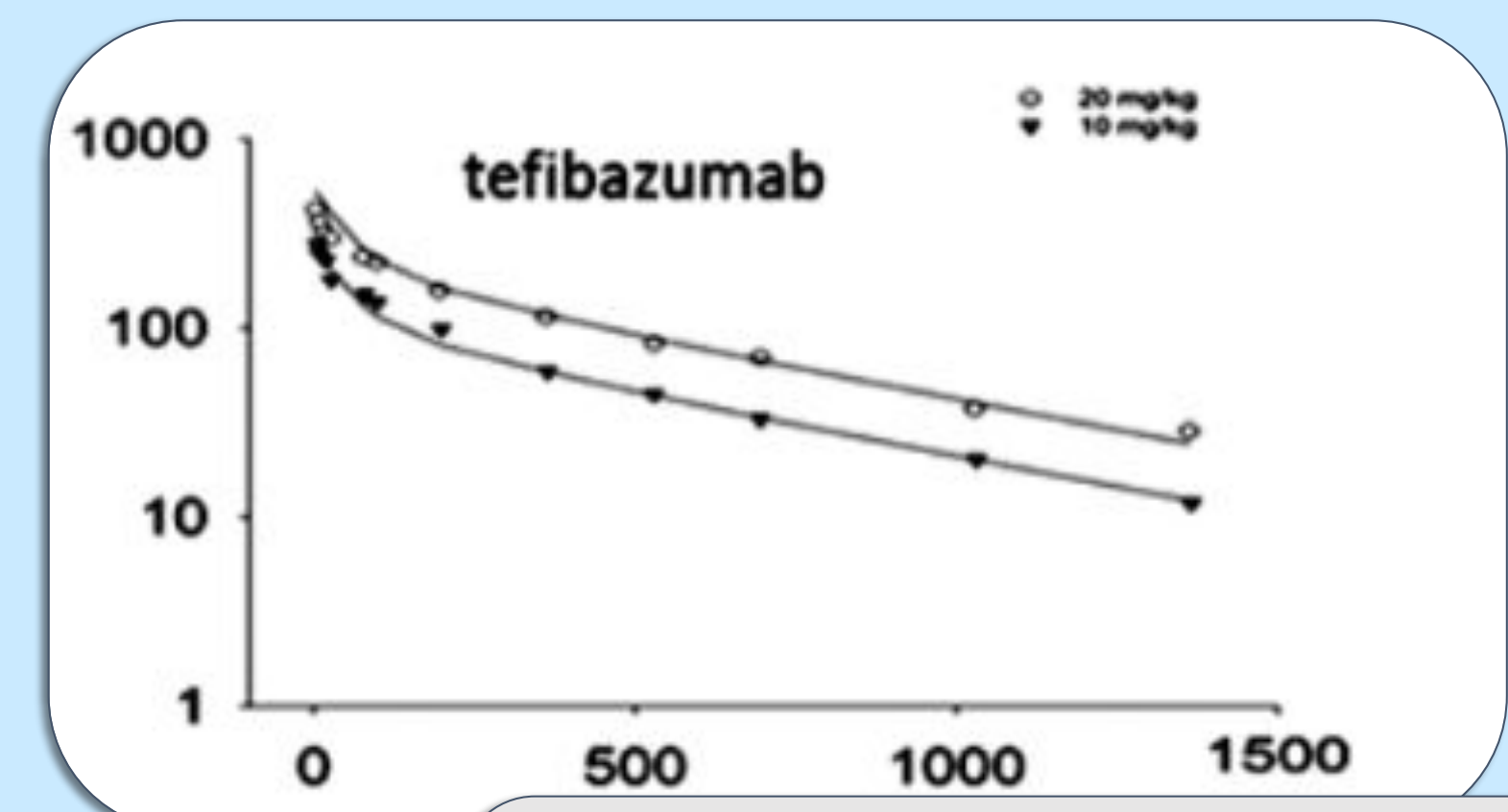
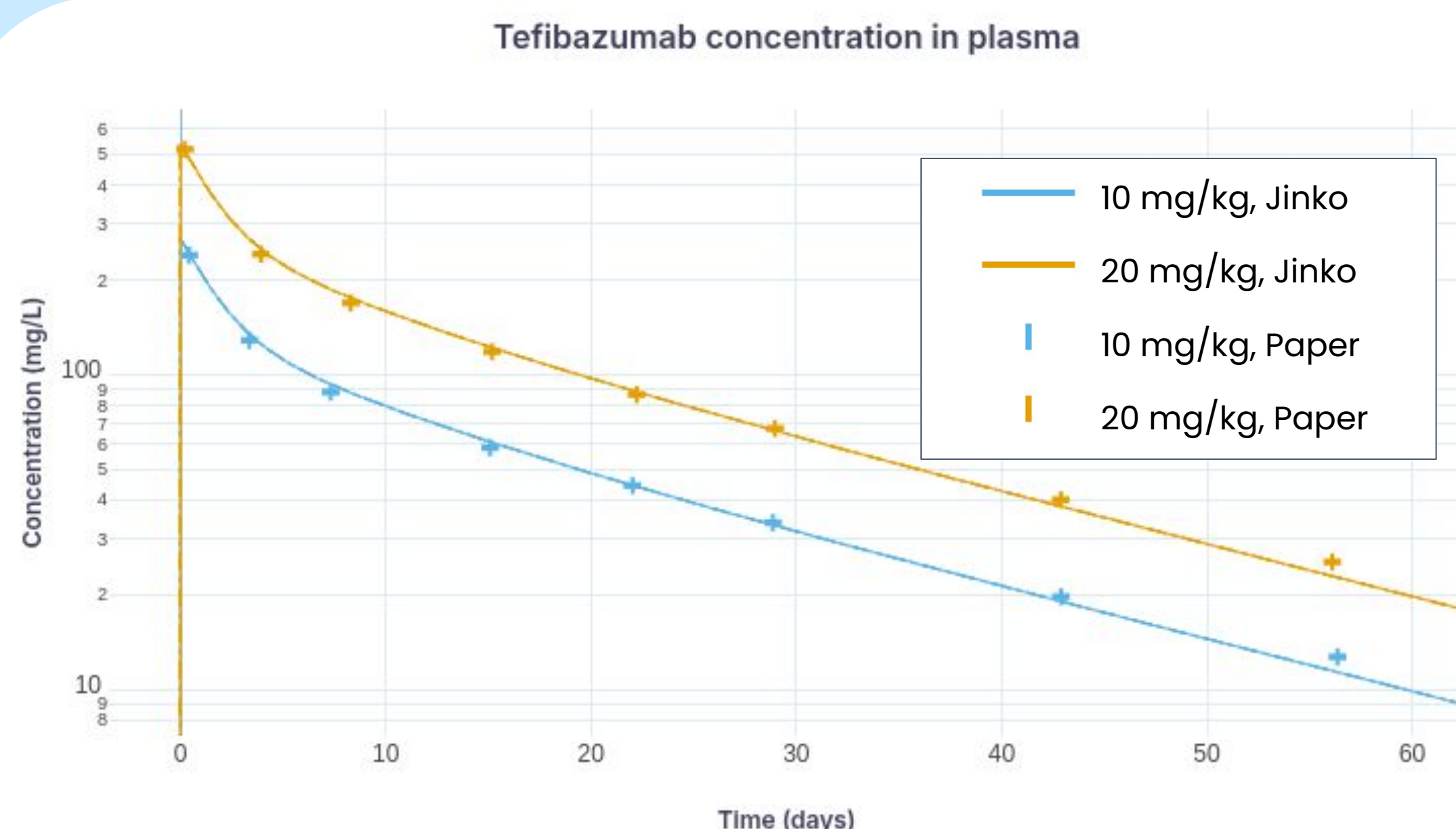
Jinkō automatic extraction

Jinkō AI conversion

Model solving

Simulation of the model and comparison to the paper outputs

- Simulation in jinkō:** Model is run on two arms for Tefibazumab (10mg/kg and 20 mg/kg).
- Model's outputs are compared to published graphs:** Extracted data points from the reference paper can be plotted alongside simulation results.



Comparison between the simulation outputs and the published graph allows to easily validate the implementation of the model

METHODS

We developed an automated pipeline that combines Optical Character Recognition (OCR) for extracting text from PDFs and scanned documents with domain-trained Natural Language Processing (NLP) to identify model components such as structural equations, parameter values, and covariate relationships.

This information is first converted into a human-readable intermediate format.

A multi-step workflow, driven by a large language model (LLM), then transforms it into a controlled, structured representation exportable to formats like SBML. The process is fully reproducible and integrated into the Jinkō platform, where models are stored in a version-controlled library with rich metadata. We evaluated the pipeline using published models.

RESULTS

The PMx models were reconstructed with high fidelity, preserving their compartments, parameter values, and variability structures. Simulated outputs from the imported models matched the published results.

CONCLUSION

- This AI-driven approach automates PMx model conversion, enhancing reproducibility, interoperability, and reuse.
- Within the Jinkō platform, these standardized models enable advanced in silico applications—from drug repositioning to virtual clinical trials—ultimately accelerating data-driven decision-making in drug development.

REFERENCES

- [1] Cucurull-Sanchez, Lourdes, et al. “Best Practices to Maximize the Use and Reuse of Quantitative and Systems Pharmacology Models: Recommendations From the United Kingdom Quantitative and Systems Pharmacology Network.” *CPT: Pharmacometrics & Systems Pharmacology*, vol. 8, no. 5, Mar. 2019, pp. 259–72.
- [2] Cao Y, Balthasar JP, Jusko WJ. Second-generation minimal physiologically-based pharmacokinetic model for monoclonal antibodies. *Journal of pharmacokinetics and pharmacodynamics*. 2013 Oct;40:597–607.