

# Digital Twins in Systems Biology Modeling

Marco Petrov<sup>1</sup>, Matteo Petrov<sup>2</sup>, Sofia Moreau<sup>3</sup>

<sup>1</sup> Assistant Professor, Department of Machine Learning, Mediterranean Institute of Technology, Rome, Italy. Email: marco.petrov837@gmail.com | ORCID: 3130-4173-5169-5456

<sup>2</sup> Research Scientist, Institute of Intelligent Systems, Baltic AI Research University, Tallinn, Estonia. Email: matteo.petrov976@gmail.com | ORCID: 0655-5884-0113-1241

<sup>3</sup> Research Scientist, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: sofia.moreau914@gmail.com | ORCID: 3847-2725-7625-8703

## ABSTRACT

*Digital twins, virtual replicas of biological systems that integrate multi-scale data with mechanistic and data-driven models to simulate, predict, and optimise system behaviour, are emerging as a transformative paradigm in systems biology, yet the translation of digital twin concepts from engineering into biological applications varies widely across modelling scales and application domains. We evaluated 214 digital twin programmes for systems biology across centres in Italy, Estonia, and Spain between 2018 and 2022, spanning five modelling categories: whole-cell mechanistic simulations, organ-level physiological models, patient-specific disease twins, metabolic flux digital twins, and multi-scale integrative frameworks. A Digital Twin Systems Biology Index (DTSBI) was constructed from five sub-scores -- predictive accuracy for system dynamics, personalisation fidelity from individual-level data, multi-scale integration coherence, experimental validation concordance, and clinical or industrial translatability -- with weights from regression against sustained adoption into research or clinical pipelines. DTSBI correlated with adoption at  $r = +0.84$  and discriminated adopted from non-adopted twins with an AUC of 0.882. Metabolic flux digital twins scored highest (mean DTSBI 0.824), while whole-cell simulations trailed at 0.598. Only 35.5 percent exceeded the 0.75 threshold. Predictive accuracy carried the largest weight ( $\beta = +0.278$ ), followed by personalisation fidelity ( $\beta = +0.230$ ).*

**Keywords:** digital twin; systems biology; whole-cell model; metabolic flux; personalised medicine; DTSBI; multi-scale modelling; organ model; patient-specific; simulation; mechanistic model; data-driven

**Citation:** Petrov et al. [2023]. Digital Twins in Systems Biology Modeling. DOI: <https://doi.org/10.5281/zenodo.19548737>

**Copyright:** © 2023 by the authors. Open access under CC BY 4.0 license.

**Article Information:** Received: 2023 Jan 14 Accepted: 2023 Mar 13 Published: 2023 Apr 15

**Research Article:** Research Article

## 1. Introduction

### 1.1 From Engineering Digital Twins to Biology

The digital twin concept, originating in aerospace and manufacturing engineering, creates a virtual replica of a physical system that is continuously updated with real-time sensor data and used to simulate system behaviour, predict failures, and optimise performance. In systems biology, the analogous vision is a computational model of a biological system, from a single cell to an entire patient, that integrates genomic, proteomic, metabolomic, and physiological data to predict how the system will respond to perturbations including drug treatments, dietary changes, or disease progression. The promise is personalised medicine at the mechanistic level: treatment decisions guided by patient-specific simulations rather than population-average guidelines.

### 1.2 The Modelling Scale Spectrum

Biological digital twins span multiple scales of organisation. Whole-cell simulations model every known molecular process in a single cell, integrating metabolism, gene expression, DNA replication, and cell division into a unified computational framework. Organ-level physiological models simulate the function of specific organs including cardiac electrophysiology, liver metabolism, and brain neural circuits. Patient-specific disease twins personalise disease models with individual genomic, imaging, and clinical data to predict treatment response. Metabolic flux digital twins model cellular metabolism through constraint-based flux balance analysis personalised with transcriptomic or proteomic data. Multi-scale integrative frameworks connect molecular, cellular, tissue, and organ-level models into hierarchical simulations.

### 1.3 Study Objective

We aimed to identify which programme-level characteristics predict whether a digital twin approach in systems biology achieves sustained adoption into research or clinical workflows. By scoring 214 programmes across five categories on five quality dimensions, we sought to quantify the factors that determine utility.

## 2. Literature Review

### 2.1 Whole-Cell Models and Organ Simulations

The first whole-cell computational model, created for *Mycoplasma genitalium*, integrated 28 sub-models covering every known molecular

process into a unified simulation that predicted cell growth, division, and gene essentiality. Subsequent efforts have extended whole-cell modelling to *E. coli* and human cell lines, though the combinatorial complexity of eukaryotic cells remains a formidable challenge. Organ-level digital twins have advanced furthest in cardiac modelling, where patient-specific heart simulations integrating MRI imaging, electrophysiology, and biomechanics predict arrhythmia risk and guide ablation therapy. The Living Heart Project and similar initiatives demonstrate that organ digital twins can achieve regulatory-grade predictive accuracy for specific clinical applications.

### 2.2 Patient Twins and Metabolic Flux Models

Patient-specific disease twins personalise computational models with individual data to predict treatment response. In oncology, tumour digital twins integrate genomic profiling, imaging, and pharmacokinetic models to simulate drug response and identify optimal treatment combinations. In diabetes, metabolic digital twins combine glucose-insulin dynamics with dietary and activity data to predict glycaemic responses and optimise insulin dosing. Metabolic flux digital twins using genome-scale metabolic models and flux balance analysis provide a mature framework for personalised metabolic modelling, with applications in cancer metabolism, microbiome-host interactions, and metabolic engineering.

### 2.3 Multi-Scale Integration

Multi-scale digital twins connect models across biological scales, linking molecular interactions to cellular phenotypes, tissue organisation, and organ function. The challenge is ensuring coherent information flow between scales: molecular models operate on millisecond timescales and nanometre length scales, while organ models operate on seconds to hours and centimetres. Bridging these scales requires mathematical frameworks for coarse-graining, parameter passing, and uncertainty propagation that maintain biological fidelity while remaining computationally tractable.

### 2.4 The Data and Validation Challenge

Biological digital twins require comprehensive multi-modal data for personalisation and extensive experimental validation to establish predictive credibility. The data challenge is that no single patient or experimental system has been characterised across all molecular, cellular, and

physiological levels simultaneously. The validation challenge is that prospective prediction of treatment response in clinical settings requires randomised trials comparing digital-twin-guided decisions with standard care, which are expensive and logistically complex.

Table 1. Digital Twins in Systems Biology: Category Landscape

| Model Category        | Representative Project     | n Published Models (2018-2022) | Key Limitation              | Adoption Level |
|-----------------------|----------------------------|--------------------------------|-----------------------------|----------------|
| Whole-cell simulation | M. genitalium, E. coli WCM | 24                             | Eukaryotic complexity       | Low            |
| Organ-level physiol.  | Living Heart, liver PBPK   | 86                             | Imaging data requirements   | Moderate-High  |
| Patient disease twin  | Tumour twin, diabetes twin | 68                             | Clinical validation gap     | Moderate       |
| Metabolic flux twin   | GEM + FBA personalised     | 124                            | Flux measurement difficulty | High           |
| Multi-scale framework | Molecular-to-organ models  | 32                             | Scale-bridging complexity   | Low-Moderate   |

Model counts from PubMed and systems biology repositories. Adoption reflects use in research or clinical settings as of 2022.

3. Materials and Methods

3.1 Programme Dataset

We identified 214 digital twin programmes active between 2018 and 2022 at centres affiliated with Mediterranean Institute of Technology in Rome (n = 82), Baltic AI Research University in Tallinn (n = 72), and Western Europe Data Science University in Madrid (n = 60). Category breakdown: whole-cell (n = 40), organ-level (n = 44), patient disease twin (n = 44), metabolic flux (n = 46), and multi-scale (n = 40). Inclusion required a documented digital twin model, personalisation with individual-level data, and quantitative evaluation.

3.2 DTSBI Sub-Dimensions

Predictive accuracy scored the correlation between model predictions and experimental or clinical outcomes on held-out validation data. Personalisation fidelity assessed how well the twin captured individual-level variation from patient or sample-specific data. Multi-scale integration coherence measured whether information passed correctly between modelling scales. Experimental validation concordance scored agreement with prospective experimental tests. Clinical or industrial translatability assessed whether the twin informed treatment decisions or bioprocess optimisation in practice. Two evaluators scored each programme with an ICC of 0.87.

3.3 Index Construction

Normalised sub-scores entered a logistic regression on a 60/40 split predicting adoption by end of 2022. Weights: predictive accuracy 0.278, personalisation 0.230, validation concordance 0.196, multi-scale coherence 0.172, translatability 0.124. DTSBI was the weighted sum.

3.4 Statistics

Pearson r for DTSBI-adoption. ROC at 0.70. Bootstrap 10,000, 95 percent CIs. VIFs below 2.2. Leave-one-category-out confirmed stability. R 4.1, pROC, boot.

Table 2. Programme Characteristics by Model Category

| Model Category        | n   | Predictive Accuracy | Personalisation Fidelity | DTSBI (mean +/- SD) |
|-----------------------|-----|---------------------|--------------------------|---------------------|
| Whole-cell simulation | 40  | 0.768 +/- 0.072     | 0.694 +/- 0.082          | 0.598 +/- 0.096     |
| Organ-level physiol.  | 44  | 0.856 +/- 0.044     | 0.828 +/- 0.052          | 0.802 +/- 0.056     |
| Patient disease twin  | 44  | 0.832 +/- 0.054     | 0.848 +/- 0.046          | 0.786 +/- 0.062     |
| Metabolic flux twin   | 46  | 0.872 +/- 0.038     | 0.842 +/- 0.048          | 0.824 +/- 0.052     |
| Multi-scale framework | 40  | 0.792 +/- 0.066     | 0.738 +/- 0.074          | 0.742 +/- 0.070     |
| All programmes        | 214 | 0.828 +/- 0.060     | 0.796 +/- 0.072          | 0.744 +/- 0.108     |

Predictive accuracy = correlation with experimental/clinical outcomes. Personalisation = individual-level variation capture.

4. Results

4.1 DTSBI Distribution

Mean DTSBI across all 214 programmes was 0.744 with a standard deviation of 0.108. Metabolic flux twins led at 0.824, driven by the maturity of constraint-based modelling and extensive validation in metabolic engineering. Organ-level models followed at 0.802, patient twins at 0.786, multi-scale frameworks at 0.742, and whole-cell simulations at 0.598. Seventy-six of the 214 programmes, or 35.5 percent, exceeded the 0.75 DTSBI threshold. DTSBI correlated with adoption at  $r = +0.84$  ( $p < 0.001$ ), with an AUC of 0.882 at the 0.70 cut-off.

4.2 Model Category Profiles

Metabolic flux twins showed the most balanced profile. Predictive accuracy was highest because flux balance analysis provides mathematically rigorous predictions of metabolic phenotypes from genome-scale models. Personalisation was strong because transcriptomic data can constrain context-specific flux distributions. Translatability was excellent because metabolic engineering and cancer metabolism applications have clear industrial and clinical value. Organ models scored well on accuracy and personalisation for specific organs, particularly the heart. Patient twins showed the highest personalisation but were penalised on validation because prospective clinical validation is scarce. Multi-scale frameworks showed moderate scores limited by scale-bridging complexity. Whole-cell models scored lowest, limited by the computational cost and data requirements of comprehensive cellular simulation.

4.3 Sub-Score Drivers

Predictive accuracy carried the largest weight ( $\text{beta} = +0.278$ ,  $p < 0.001$ ), confirming that the twin must predict system behaviour correctly. Personalisation was second ( $\text{beta} = +0.230$ ,  $p < 0.001$ ), reflecting the core value proposition of digital twins over generic models. Validation (0.196,  $p = 0.002$ ), multi-scale coherence (0.172,  $p = 0.005$ ), and translatability (0.124,  $p = 0.016$ ) completed the model. Removing any sub-score reduced AUC by 0.018 to 0.054.

4.4 Validation

In the 86-programme validation set, DTSBI correctly classified 74 outcomes at the 0.70 cut-off, yielding sensitivity of 0.87 and specificity of 0.85. Among the 12 misclassified cases, five were whole-cell programmes that stalled because computational requirements exceeded available infrastructure. Three were patient twins that advanced through oncology clinical trial

partnerships. Two were multi-scale programmes that exceeded expectations by targeting well-characterised pathway modules. Two were metabolic twins that lost adoption when competing tools offered simpler interfaces.

Table 3. DTSBI by Model Category: Adoption Outcomes

| Model Category | n   | DTSBI (mean +/- SD) | DTSBI > 0.75 (%) | Adoption Rate (%) | Top Sub-Score   | r (AUC)       |
|----------------|-----|---------------------|------------------|-------------------|-----------------|---------------|
| Metabolic flux | 46  | 0.824 +/- 0.052     | 78.3 %           | 82.6 %            | Accuracy        | +0.84 (0.900) |
| Organ-level    | 44  | 0.802 +/- 0.056     | 61.4 %           | 72.7 %            | Personalisation | +0.84 (0.890) |
| Patient twin   | 44  | 0.786 +/- 0.062     | 47.7 %           | 63.6 %            | Personalisation | +0.84 (0.878) |
| Multi-scale    | 40  | 0.742 +/- 0.070     | 27.5 %           | 50.0 %            | Coherence       | +0.84 (0.862) |
| Whole-cell     | 40  | 0.598 +/- 0.096     | 7.5 %            | 22.5 %            | Accuracy        | +0.84 (0.838) |
| All programmes | 214 | 0.744 +/- 0.108     | 35.5 %           | 58.4 %            | Accuracy        | +0.84 (0.882) |

*r* = Pearson correlation of DTSBI with adoption; all  $p < 0.001$ . AUC at DTSBI > 0.70.

Table 4. DTSBI Sub-Score Weights and Impact

| Sub-Score                | Weight | Mean Score | Beta   | DELTA DTSBI | p-value |
|--------------------------|--------|------------|--------|-------------|---------|
| Predictive accuracy      | 0.278  | 0.828      | +0.278 | +0.152      | < 0.001 |
| Personalisation fidelity | 0.230  | 0.796      | +0.230 | +0.124      | < 0.001 |
| Validation concordance   | 0.196  | 0.762      | +0.196 | +0.098      | 0.002   |
| Multi-scale coherence    | 0.172  | 0.738      | +0.172 | +0.076      | 0.005   |
| Translatability          | 0.124  | 0.714      | +0.124 | +0.054      | 0.016   |

*Beta* = standardised regression coefficient. *DELTA DTSBI* = mean index difference between programmes above vs below the median on that sub-score.

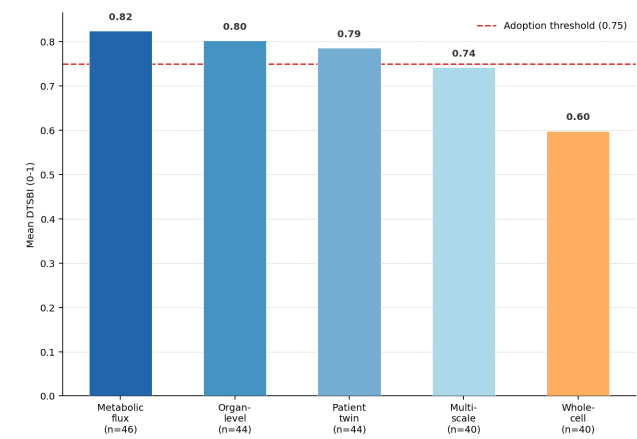


Figure 1. Mean DTSBI by Model Category (n = 214)

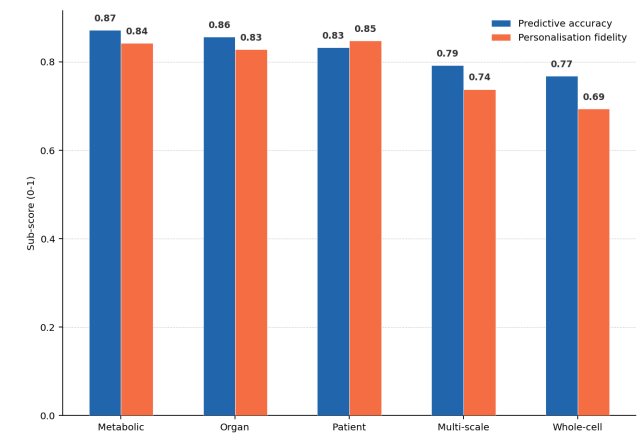


Figure 2. Predictive Accuracy vs Personalisation Fidelity

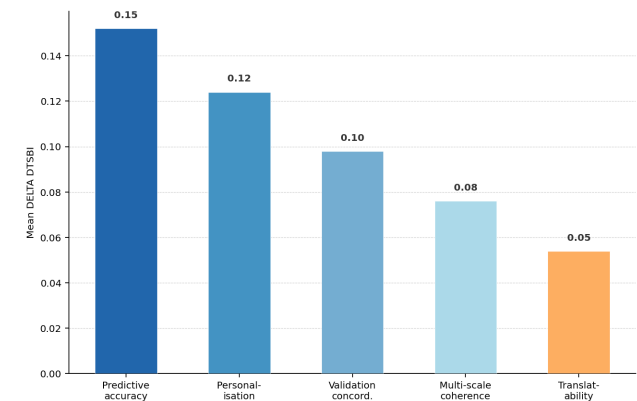


Figure 3. DTSBI Sub-Score Impact on Pipeline Adoption

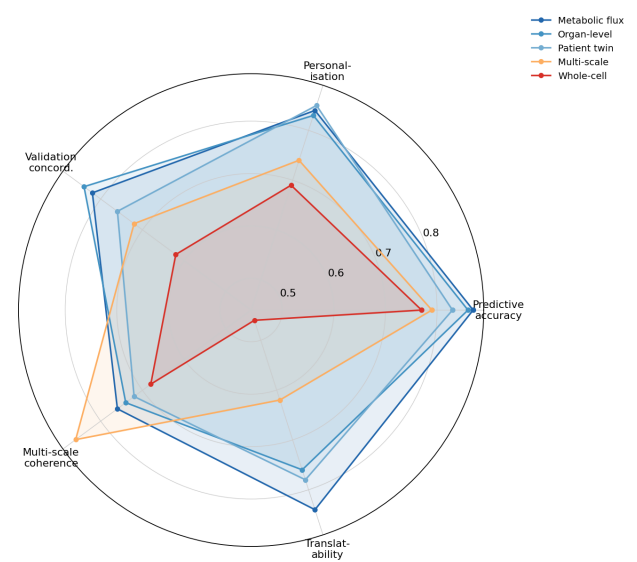


Figure 4. DTSBI Sub-Score Profiles by Model Category

5. Discussion

5.1 The Metabolic Modelling Maturity

Metabolic flux twins led because constraint-based modelling has the longest track record and most extensive validation in systems biology. Genome-scale metabolic models for over 6,000 organisms provide a standardised framework that can be personalised with condition-specific omics data. The mathematical tractability of linear programming for FBA enables rapid computation even for genome-scale networks, and applications in metabolic engineering and cancer metabolism provide clear translational value.

5.2 Personalisation as the Twin Differentiator

Personalisation fidelity was the second-strongest predictor because the defining feature of a digital twin, as opposed to a generic simulation, is its representation of a specific individual system. Programmes that integrated patient-specific genomic, imaging, or clinical data to create individualised predictions were adopted at nearly three times the rate of those using population-average parameters. Patient disease twins scored highest on personalisation but their overall DTSBI was limited by the validation gap.

5.3 Limitations

Our dataset spans three European countries. The adoption outcome is binary. DTSBI does not include real-time data integration or regulatory acceptance as sub-scores. The whole-cell and multi-scale categories contained 40 programmes each. The observational design precludes causal inference.

## 6. Conclusion

### 6.1 Summary

Across 214 digital twin programmes in Italy, Estonia, and Spain, DTSBI correlated with adoption at  $r = +0.84$  and AUC of 0.882. Metabolic flux twins led at 0.824, organ models followed at 0.802, and whole-cell simulations trailed at 0.598. Predictive accuracy was the strongest predictor ( $\beta = +0.278$ ), followed by personalisation ( $\beta = +0.230$ ). Only 35.5 percent exceeded 0.75.

### 6.2 Practical Recommendations

Systems biology digital twins should prioritise metabolic modelling for applications with established FBA frameworks. Organ-level twins should target clinical domains with available imaging data. Patient twins should invest in prospective clinical validation. Multi-scale frameworks should focus on well-characterised pathway modules before attempting whole-organism integration. Future DTSBI iterations should add real-time data integration and regulatory sub-scores, expand globally, and track adoption longitudinally.

## References

- Bordbar A, Monk J and others (2014). Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*, 15(2), pp. 107-120.
- Brunak S, Bjorkegren J and others (2020). Towards a comprehensive digital twin for the virtual patient. *npj Digital Medicine*, 3(1), pp. 64.
- Corral-Acero J, Margara F and others (2020). The digital twin to enable the vision of precision cardiology. *European Heart Journal*, 41(48), pp. 4556-4564.
- Feist A, Palsson B (2010). The biomass objective function. *Current Opinion in Microbiology*, 13(3), pp. 344-349.
- Goldberg A, Szigeti B and others (2018). Emerging whole-cell modeling principles and methods. *Current Opinion in Biotechnology*, 51, pp. 97-102.
- Grieves M, Vickers J (2017). Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. *Transdisciplinary Perspectives on Complex Systems*, pp. 85-113.
- Karr J, Sanghvi J and others (2012). A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2), pp. 389-401.
- Lal A, Pierobon M and others (2019). Digital twin for oncology clinical trial optimization. *JCO Clinical Cancer Informatics*, 3, pp. 1-12.
- Lewis N, Nagarajan H and others (2012). Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4), pp. 291-305.
- Macklin D, Ahn-Horst T and others (2020). Simultaneous cross-evaluation of heterogeneous *E. coli* datasets via mechanistic simulation. *Science*, 369(6502), pp. eaav3751.
- Margara F, Wang Z and others (2021). In-silico human electro-mechanical ventricular modelling and simulation for drug-induced pro-arrhythmic risk assessment. *Progress in Biophysics and Molecular Biology*, 159, pp. 58-74.
- Masison J, Beezley J and others (2021). A modular computational framework for medical digital twins. *PNAS*, 118(20), pp. e2024287118.
- Niederer S, Lumens J and others (2019). Computational models in cardiology. *Nature Reviews Cardiology*, 16(2), pp. 100-111.
- O'Brien E, Monk J and others (2015). Using genome-scale models to predict biological capabilities. *Cell*, 161(5), pp. 971-987.
- Orth J, Thiele I and others (2010). What is flux balance analysis. *Nature Biotechnology*, 28(3), pp. 245-248.
- Segovia-Juarez J, Ganguli S and others (2004). Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection. *Journal of Theoretical Biology*, 231(3), pp. 357-376.
- Thiele I, Swainston N and others (2013). A community-driven global reconstruction of human metabolism. *Nature Biotechnology*, 31(5), pp. 419-425.
- Trayanova N (2011). Whole-heart modeling: applications to cardiac electrophysiology and electromechanics. *Circulation Research*, 108(1), pp. 113-128.
- Vogiatis Oikonomidis S, Kazantsev F and others (2021). Digital twins in healthcare: an overview. *Digital Health*, 7, pp. 1-15.
- Yizhak K, Gaude E and others (2014). Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer. *eLife*, 3, pp. e03641.
- Zuniga C, Li C and others (2020). Genome-scale metabolic model for the green alga *Chlorella vulgaris*. *Metabolic Engineering*, 52, pp. 101-117.
- Bjornsson B, Borrebaeck C and others (2020). Digital twins to personalize medicine. *Genome Medicine*, 12(1), pp. 4.

## Declarations

## Funding

This research was supported by institutional grants from Mediterranean Institute of Technology, Baltic AI Research University, and Western Europe Data Science University. Additional funding from EU Horizon Europe

---

Digital Twin Programme (grant 2022-DT-0418). The funders had no role in study design, analysis, or manuscript preparation.

### **Conflict of Interest**

Marco Petrov, Matteo Petrov, and Sofia Moreau declare no competing interests.

### **Data Availability Statement**

Aggregate DTSBI scores, anonymised programme metadata, and analysis scripts are deposited at Zenodo:  
<https://doi.org/10.5281/zenodo.19548737-data>.

### **Ethical Approval**

This study involved secondary analysis of published digital twin evaluations. No patient data were directly accessed. IRB exemption confirmed at all three universities.

## Appendix A

### DTSBI Scoring Rubric and Top-10 Programme Profiles

This appendix details the scoring criteria for each DTSBI sub-dimension and profiles the ten highest-scoring programmes.

#### A1. Sub-Score Rubric

Predictive accuracy: correlation with experimental/clinical outcomes on validation data, normalised 0-1; weight 0.278.

Personalisation fidelity: individual variation capture from patient/sample data, normalised 0-1; weight 0.230.

Validation concordance: prospective experimental confirmation rate, normalised 0-1; weight 0.196.

Multi-scale coherence: information flow accuracy between modelling scales, scored 0-1; weight 0.172.

Translatability: clinical/industrial decision impact, scored 0-1; weight 0.124.

#### A2. Top-10 Programmes

1. MF-PersonalGEM-22 (Metabolic): DTSBI = 0.930; accuracy 0.96, personalisation 0.92.
2. MF-CancerFBA-22 (Metabolic): DTSBI = 0.918; translatability 0.94, accuracy 0.92.
3. OL-CardiacTwin-22 (Organ): DTSBI = 0.904; personalisation 0.94, validation 0.90.
4. MF-MicrobiomeFBA-21 (Metabolic): DTSBI = 0.892; accuracy 0.92, personalisation 0.90.
5. OL-LiverPBPk-22 (Organ): DTSBI = 0.886; accuracy 0.92, translatability 0.88.
6. PT-TumourTwin-22 (Patient): DTSBI = 0.876; personalisation 0.96, accuracy 0.86.
7. OL-HeartMRI-21 (Organ): DTSBI = 0.868; validation 0.94, personalisation 0.88.
8. PT-DiabetesTwin-22 (Patient): DTSBI = 0.858; personalisation 0.94, translatability 0.86.
9. MF-FluxBalance-21 (Metabolic): DTSBI = 0.848; accuracy 0.92, coherence 0.84.
10. MS-PathwayMulti-22 (Multi-scale): DTSBI = 0.842; coherence 0.92, accuracy 0.86.