

Language Driven Therapy Design in Predictive Oncology

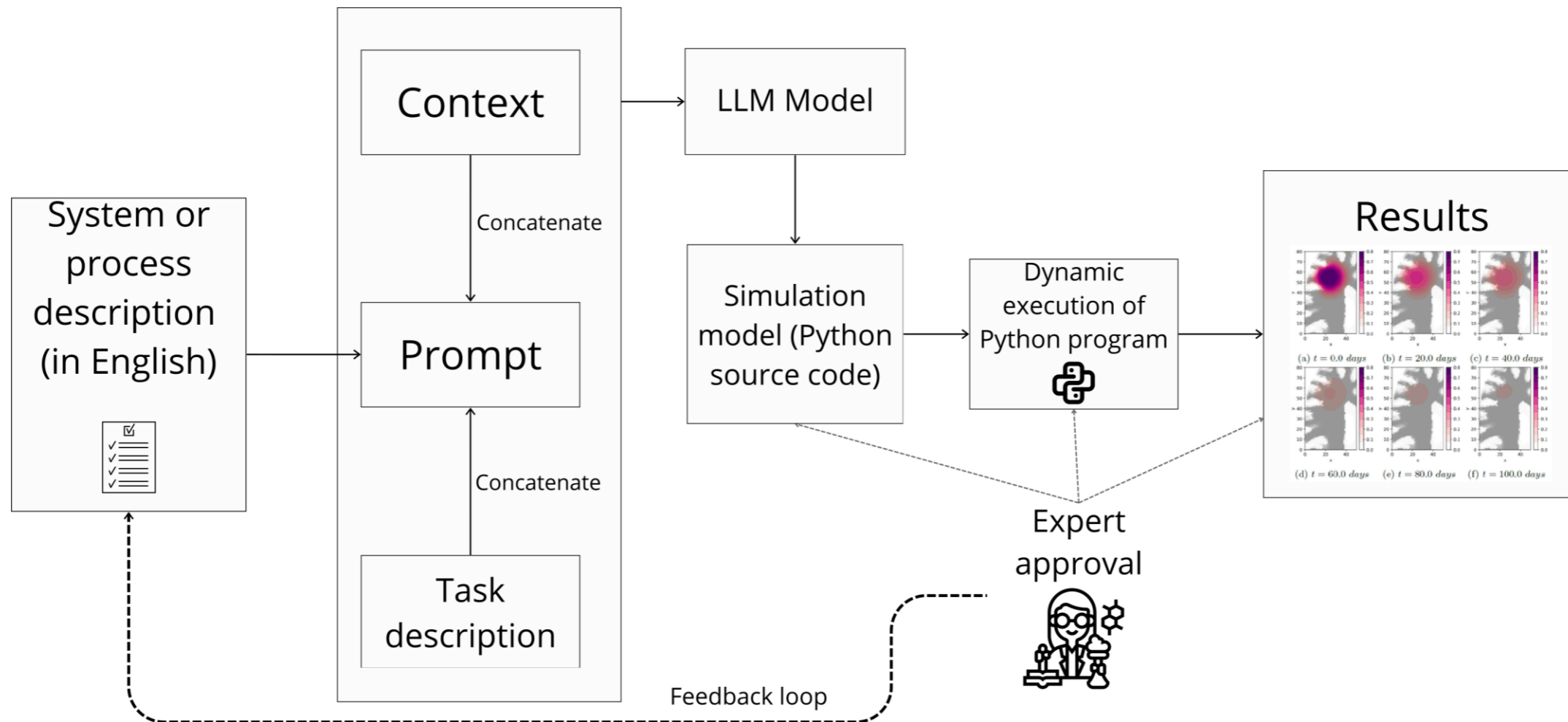
14 April 2026

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
AGH University of Krakow

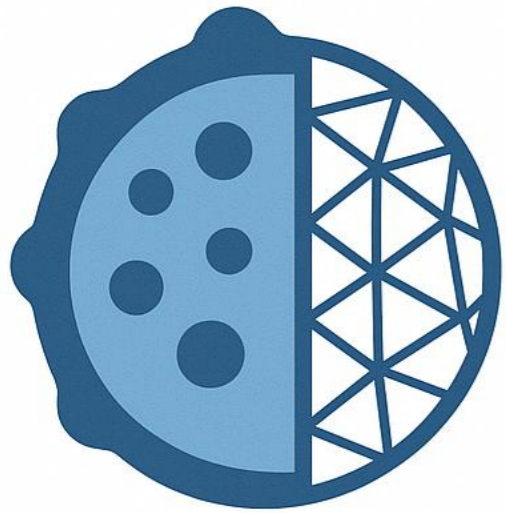


What problem are we addressing?

Modern oncology remains constrained by a persistent methodological divide: biological and clinical knowledge is predominantly articulated in narrative form, while predictive treatment planning requires formal, executable models capable of simulating tumor evolution under therapeutic pressure. Here, we argue that the emergence of Large Language Models (LLMs) enables a fundamental reconfiguration of this paradigm. Their capability motivates the concept of conversational modeling, in which modeling pipelines are initiated, shaped, and iteratively refined through natural language interaction rather than explicit mathematical programming.



Comparison

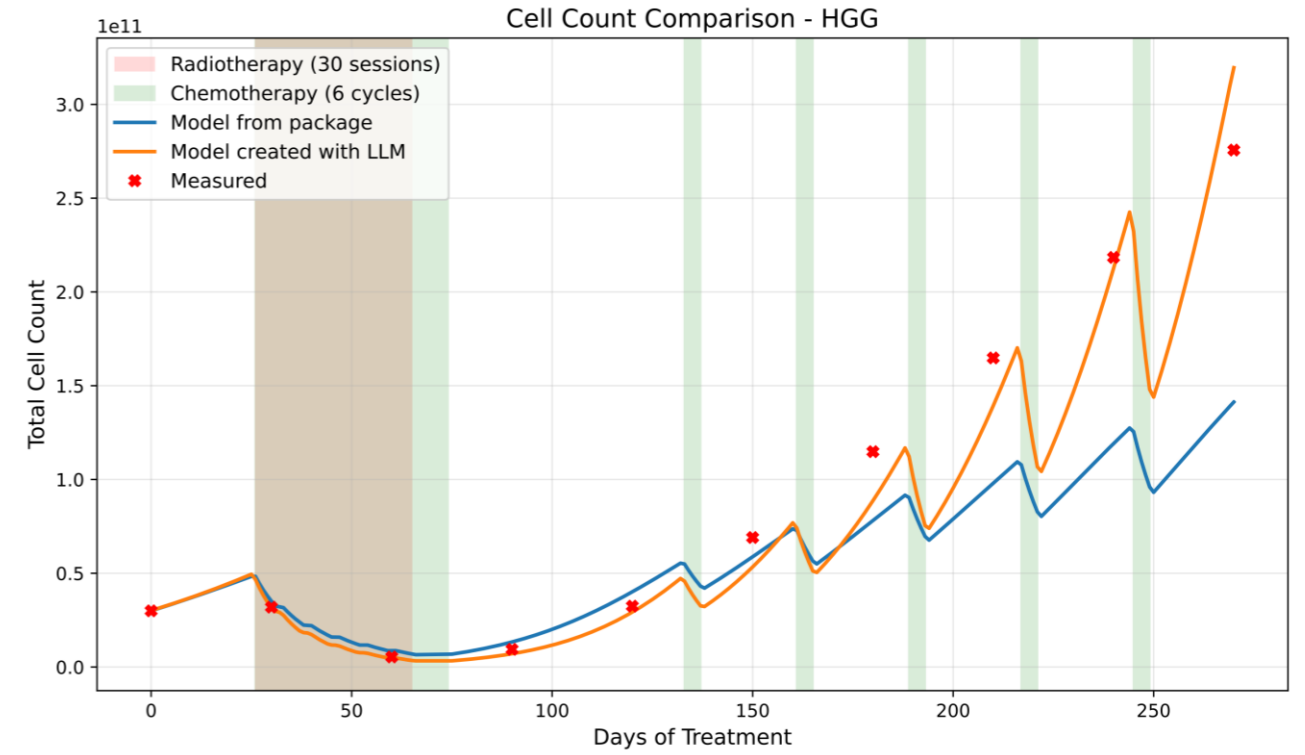
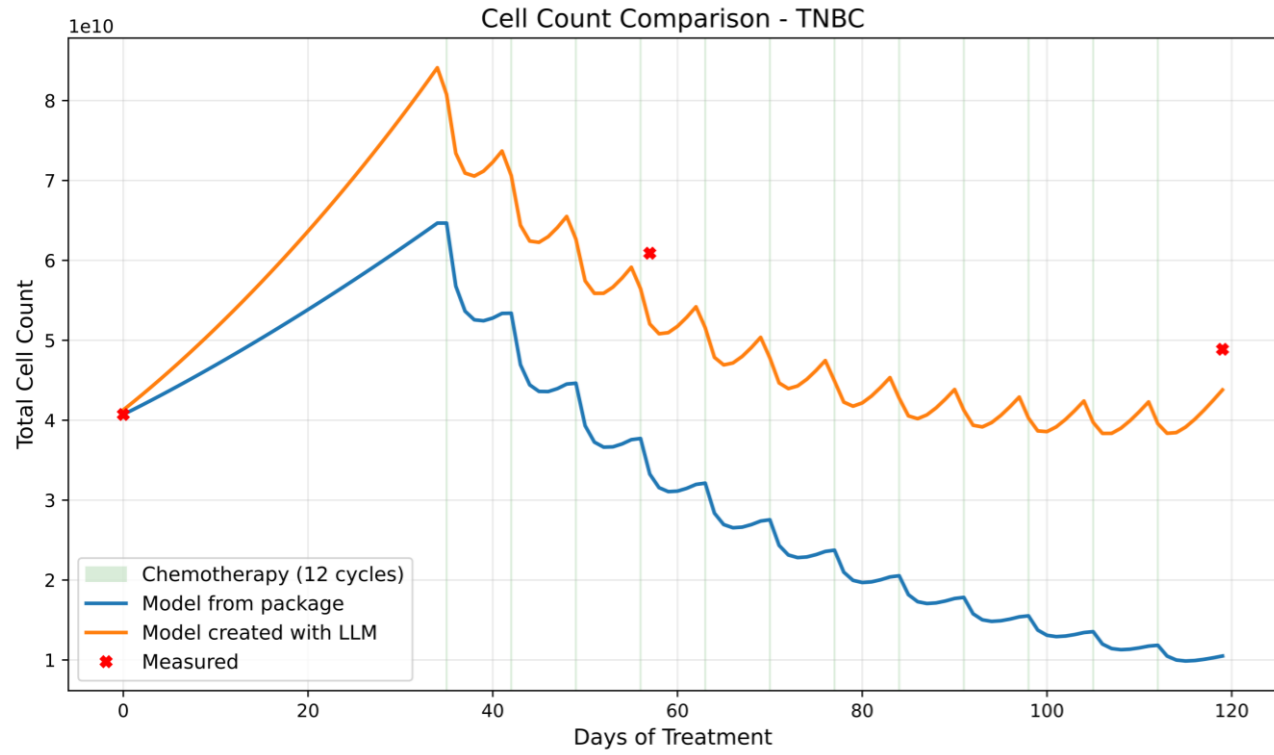


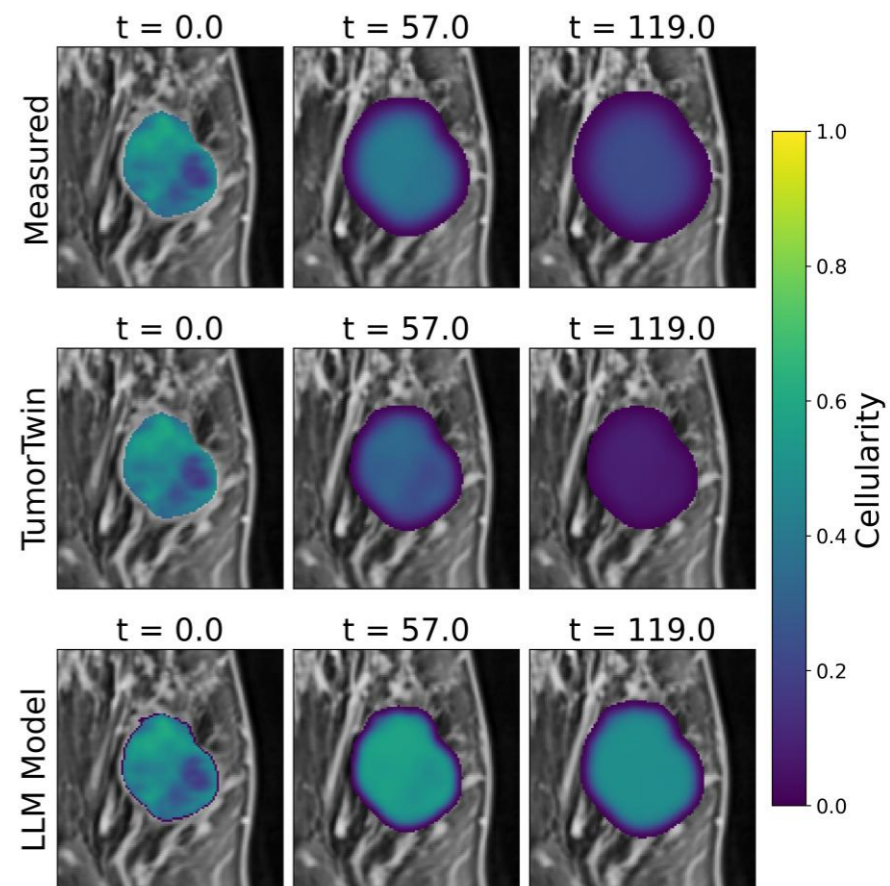
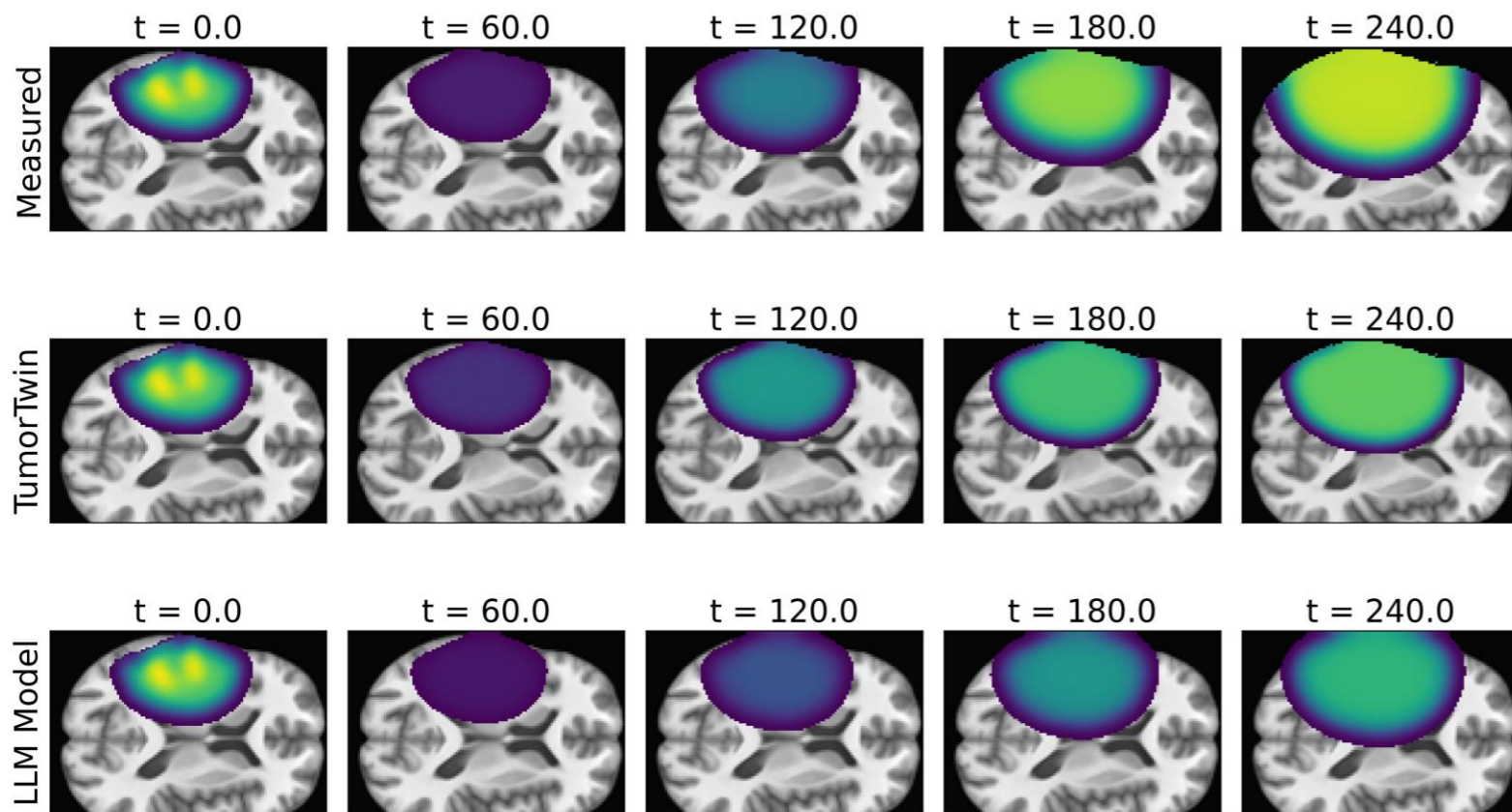
Tumor Twin

Kapteyn, M., et al.: Tumortwin: A python framework for patient-specific digital twins in oncology (2025)

$$\frac{\partial N}{\partial t} = \underbrace{\nabla \cdot (D \nabla N)}_{\text{invasion}} + \underbrace{k(\mathbf{x}) N \left(1 - \frac{N}{\theta}\right)}_{\text{logistic growth}} - \underbrace{N(\mathbf{x}, t) \sum_{i=1}^{n_{CT}} \sum_{j=1}^{T_i} \alpha_i C_i e^{-\beta_i (t - \tau_{i,j})}}_{\text{chemotherapy}}$$

$$N(\mathbf{x}, t^+) = \underbrace{N(\mathbf{x}, t^-) \exp \left[-\alpha_{RT} d_{RT}(t) - \beta_{RT} d_{RT}(t)^2 \right]}_{\text{radiotherapy}}$$





Cell count metrics

Metric	TumorTwin HGG	LLM HGG	TumorTwin TNBC	LLM TNBC
MSE	$3.38 * 10^{21}$	$3.55 * 10^{20}$	$7.46 * 10^{20}$	$3.48 * 10^{19}$
RMSE	$5.82 * 10^{10}$	$1.88 * 10^{10}$	$2.73 * 10^{10}$	$5.90 * 10^9$
MAE	$3.65 * 10^{10}$	$1.23 * 10^{10}$	$2.2 * 10^{10}$	$4.83 * 10^9$
MAPE	31.98%	12.26%	41.3%	8.77%

Dice coefficient

Visit HGG	1	2	3	4	5	6	7	8	9	10
TumorTwin	1	0.9515	0.9857	0.9480	0.8935	0.8448	0.8024	0.7712	0.7481	0.7325
LLM Model	1	0.9513	0.9725	0.9521	0.9002	0.9123	0.8823	0.8445	0.8482	0.8151

Visit TNBC	1	2	3
TumorTwin	1	0.7785	0.6043
LLM Model	1	0.7612	0.8732

Biological Narrative

Describe the cancer scenario: tumor type, growth characteristics, treatment protocol...

LLM Provider

vllm

1. Analyze Biology

2. Generate Equations

Modify Equations

Modification Request

e.g., Use Gompertz growth instead of logistic...

Refine Equations

3. Generate Code

</> Generated Python Code

</>

Model Name

Leave blank to use class name

Save Model

Saved Models

model1

Refresh

Model Source Code

```
1 import numpy as np
2 from typing import Dict, List, Tuple
3 from ldt.models.base import TumorModel, ModelParameters
4
5
6 class GliomaReactionDiffusionModel(TumorModel):
7     """
14
15     def compute_proliferation(self, c: np.ndarray) -> np.ndarray: _
28
29     def compute_treatment_effect(
30         self,
31         c: np.ndarray,
32         t: float,
33         current_day: int,
34         rt_schedule: Dict[int, float],
35         chemo_schedule: List[Tuple[int, int, float]],
36         current_drug_conc: float,
37     ) -> Tuple[np.ndarray, float]: _
77
78     def get_model_description(self) -> str: _
```

Fix Model with LLM

Patient Data

Select Patient

HGG_demo_001

Load Patient Data

Refresh Patients

Patient: [HGG_demo_001](#) (HGG) **Visits:** 10 | **Initial condition shape:** (77, 116, 36) **Radiotherapy:** 30 sessions **Chemotherapy:** 6 cycle(s) (74 entries)

Initial condition loaded from Visit 1 (ADC + cellularity). Treatment schedule relative to Visit 1 (day 0).

Run Simulation

Simulation Days

279

1  365

Model Parameters (JSON)

```
{
  "D": 0.025,
  "rho": 0.05,
  "K": 1.0,
  "alpha": 0.035,
  "beta": 0.0035,
  "decay_rate": 9.24
}
```

Tumor Burden Over Time



Run Simulation

Usage of Cyfronet infrastructure

- Self hosted LLMs using inference engines like vLLM on NVIDIA A100 GPUs on Athena supercomputer
- Running models on node's CPU (GPU accelerated models in the future)
- (future) – storing and accessing datasets

Problems

- Not enough data to validate models
- What metrics should be used? Is cell count and Dice similarity coefficient enough?
- Is elimination of the tumor sole factor in successful treatment? What about negative effects of e.g. radiation?

Acknowledgements

This work was supported by the National Science Centre (NCN), Poland, under grant OPUS-29, **DEC-2025/57/B/ST6/04377** and PLGrid Cyfronet grant **PLG/2025/018971**.