

Agents, digital twins, synthetic arms

Julien Vibert, MD, PhD

Drug Development Department (DITEP) - Sarcoma Team

INSERM U1361 - Computational Oncology

PRISM - National PReclSion Medicine Center in Oncology

Gustave Roussy

Postdoc @ Stefan Fröhling Team / Jakob Kather Lab

NCT Heidelberg, DKFZ, University Heidelberg

13th Belgian Multidisciplinary Meeting on Urological Cancers (BMUC)
2026, Brussels

Conflicts of interest

- I have no potential conflict of interest to report



Key definitions

Artificial Intelligence (AI)

Any technique enabling machines to mimic human cognition. In oncology: from simple decision trees to complex multi-modal models.

Machine Learning (ML)

Algorithms that improve through experience, learning patterns from data instead of hand-coded rules. E.g. treatment response prediction, survival models.

Deep Learning (DL)

Neural networks with multiple layers that excel at images, sequences, and unstructured data. Powers pathology analysis, radiology, and genomics.

Generative AI

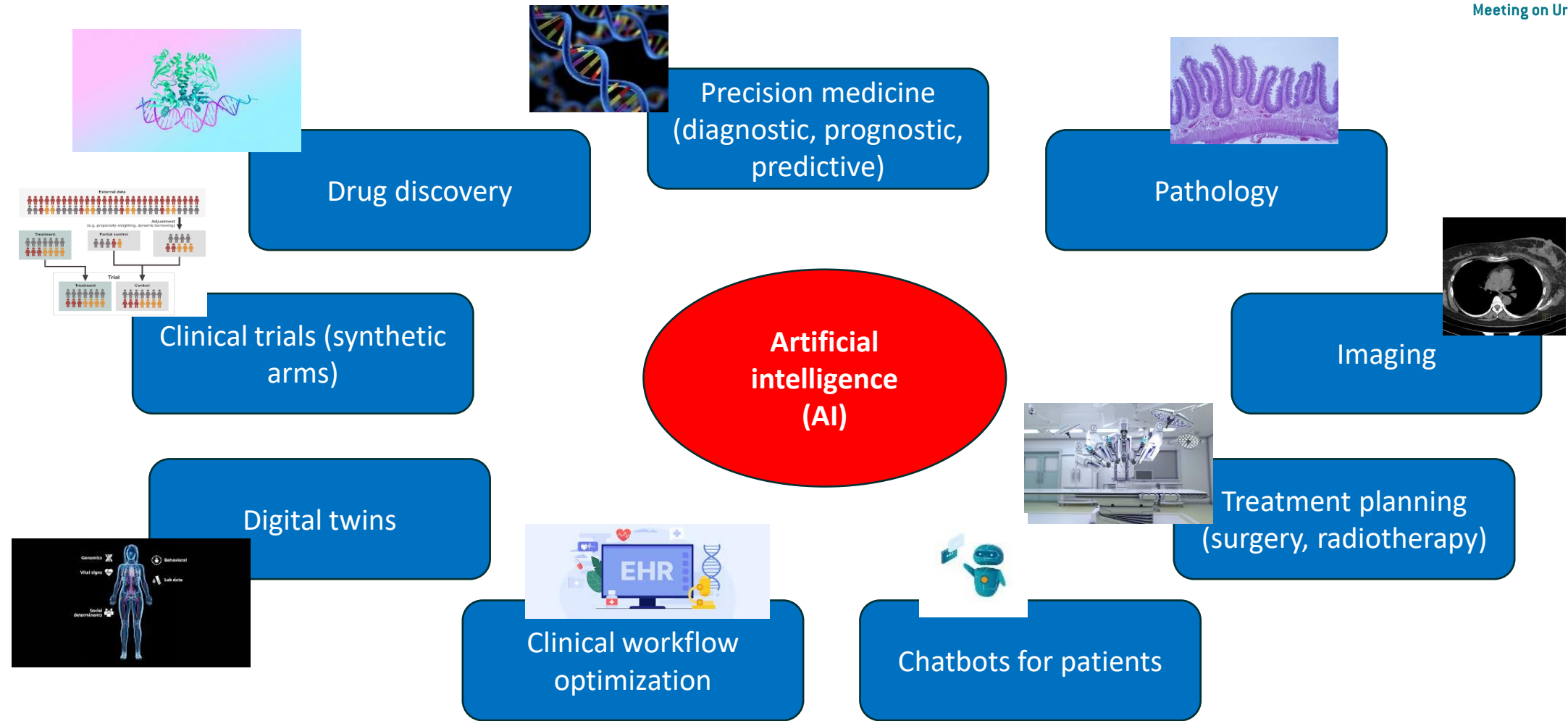
Models that generate new content rather than just classify. Produce text, images, molecular structures, and synthetic clinical data. Key architectures: GANs, VAEs, diffusion models, transformers.

Large Language Models (LLMs)

Trained on billions of text tokens to understand and produce language. Can summarize records, draft reports, and support clinical decisions. Examples: GPT-4, Claude, Gemini, Med-PaLM.



AI in Oncology



Why this matters now in GU oncology

Three forces are converging in oncology research.

1. Multimodal cancer data are exploding

Pathology, CT/MRI, genomics, PSA/ctDNA, longitudinal labs, and outcomes are increasingly linkable.

2. Foundation models changed the tooling

Generative models and LLMs do not only predict - they retrieve, summarize, create, reason, and increasingly act through tools.

3. Clinical research and care need speed and flexibility

Precision oncology fragments populations, standard of care evolves fast, and resistance is dynamic. Static workflows are increasingly mismatched.

The key shift: AI is becoming an infrastructure layer for cancer clinical research and care, not just another isolated prediction model.

Four important concepts

Predictive model

Uses past examples to estimate a future outcome: diagnosis, prognosis, response, toxicity, eligibility...

Generative model

Creates new content: molecules, images, text, trial summaries, synthetic records, candidate designs...

AI agent

A model that can reason over multiple steps, use tools, retrieve new information, and act on a workflow.

Digital twin

A dynamic patient-level virtual representation updated over time and used for prediction or simulation.

Common confusion to avoid: synthetic data are not digital twins, and chatbots are not necessarily agents.

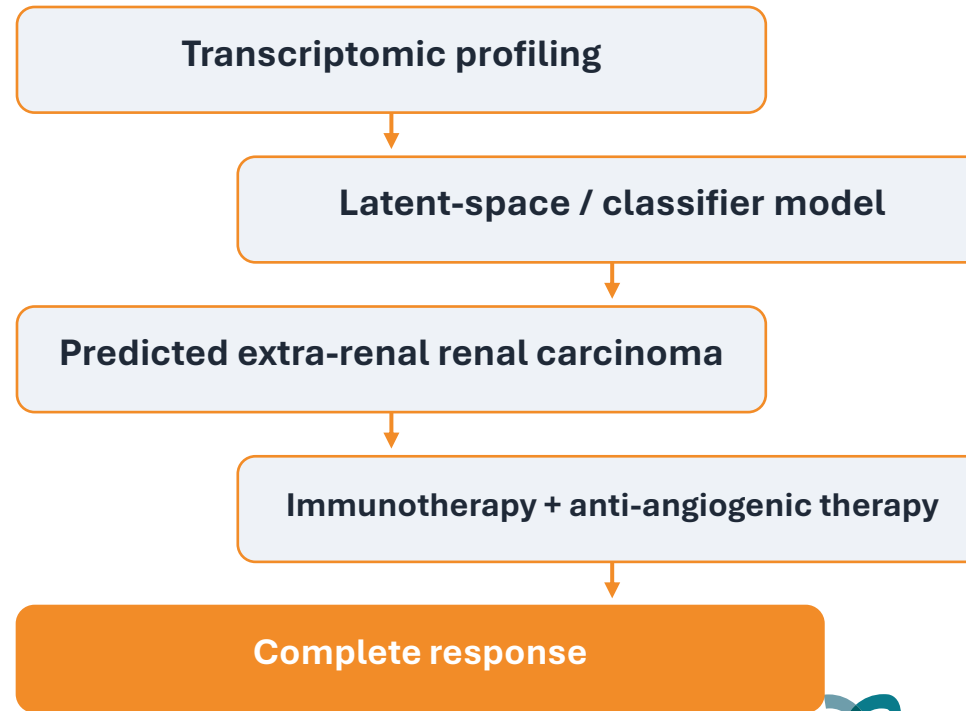
UROLOGY - RADIATION ONCOLOGY - MEDICAL ONCOLOGY

**13th Belgian Multidisciplinary
Meeting on Urological Cancers**

A motivating case: when AI changed the therapeutic hypothesis

Young patient with carcinoma of unknown primary

- 30-year-old patient
- Rapid clinical deterioration, abdominal mass, diffuse bone lesions
- Referred for suspected bone sarcoma
- Pathology: unclassified carcinoma
- Initial treatment hypothesis: pancreatic-type regimen?



Why open with this case? Because AI mattered by reframing the disease - not by replacing the oncologist.

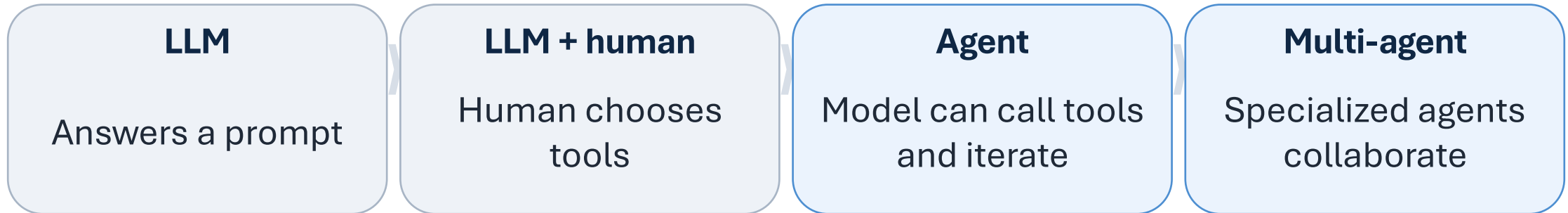
Where AI can act across cancer clinical research



The key message of today's talk: AI is most interesting when it links these stages together, not when it optimizes only one isolated prediction task.

- Upstream: find better targets, biomarkers, doses. E.g. RCC state biology, FGFR3 in bladder.
- Operationally: match patients to trials, adapt protocols. E.g. IO/TKI sequencing in mRCC.
- Longitudinally: model response and resistance. E.g. ctDNA/PSA monitoring, treatment adaptation.

From LLMs to agents: when a model can use tools

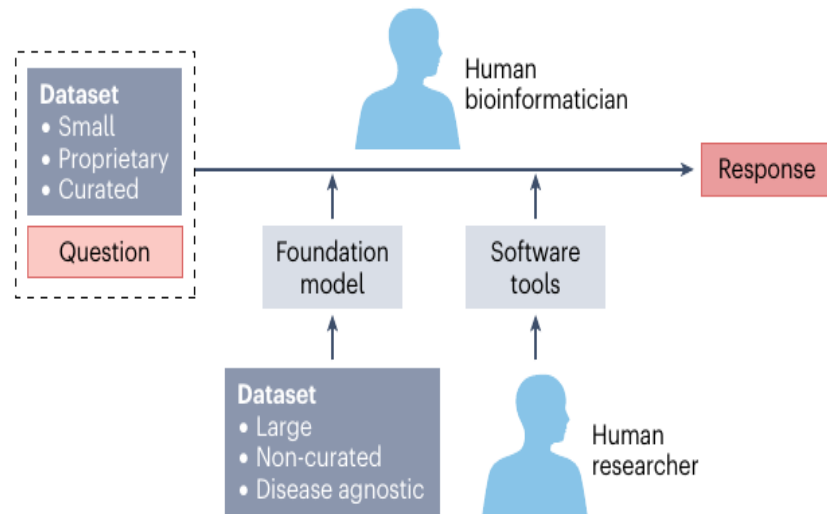


An agent is not just a chatbot. It can retrieve information, call software tools, inspect outputs, and iterate over multistep tasks.

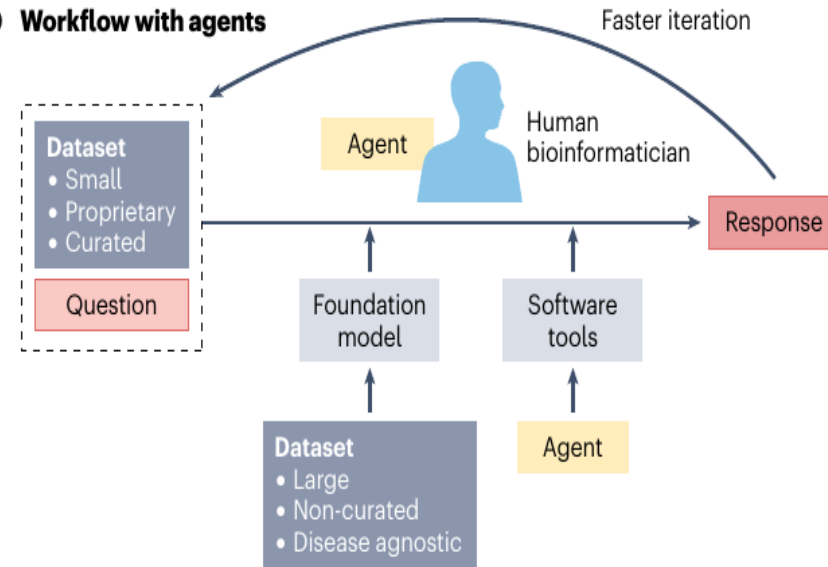
- That matters in oncology because useful work is rarely one-step text generation.
- Research workflows require databases, guidelines, pathology/radiology tools, protocol documents, and often repeated reasoning.

A useful mental model: agents reduce the friction of tool-heavy workflows

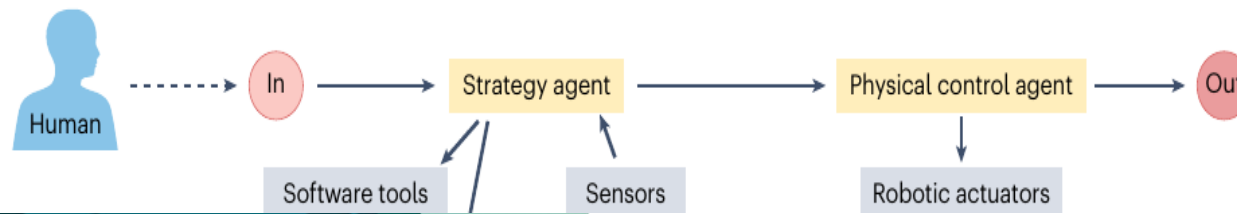
a Human bioinformatics workflow



b Workflow with agents



c Agent-based physical laboratory



Where agents could help in oncology research workflows

Literature and guideline synthesis

Find, compare, and trace evidence faster.

Protocol drafting and amendment support

Summarize rationale, criteria, endpoints, and design alternatives.

Molecular report interpretation

Connect variants and signatures to knowledge bases and therapies.

Trial matching and pre-screening

Reason over unstructured charts and protocol criteria.

Tumour board prep

Create a structured dossier.

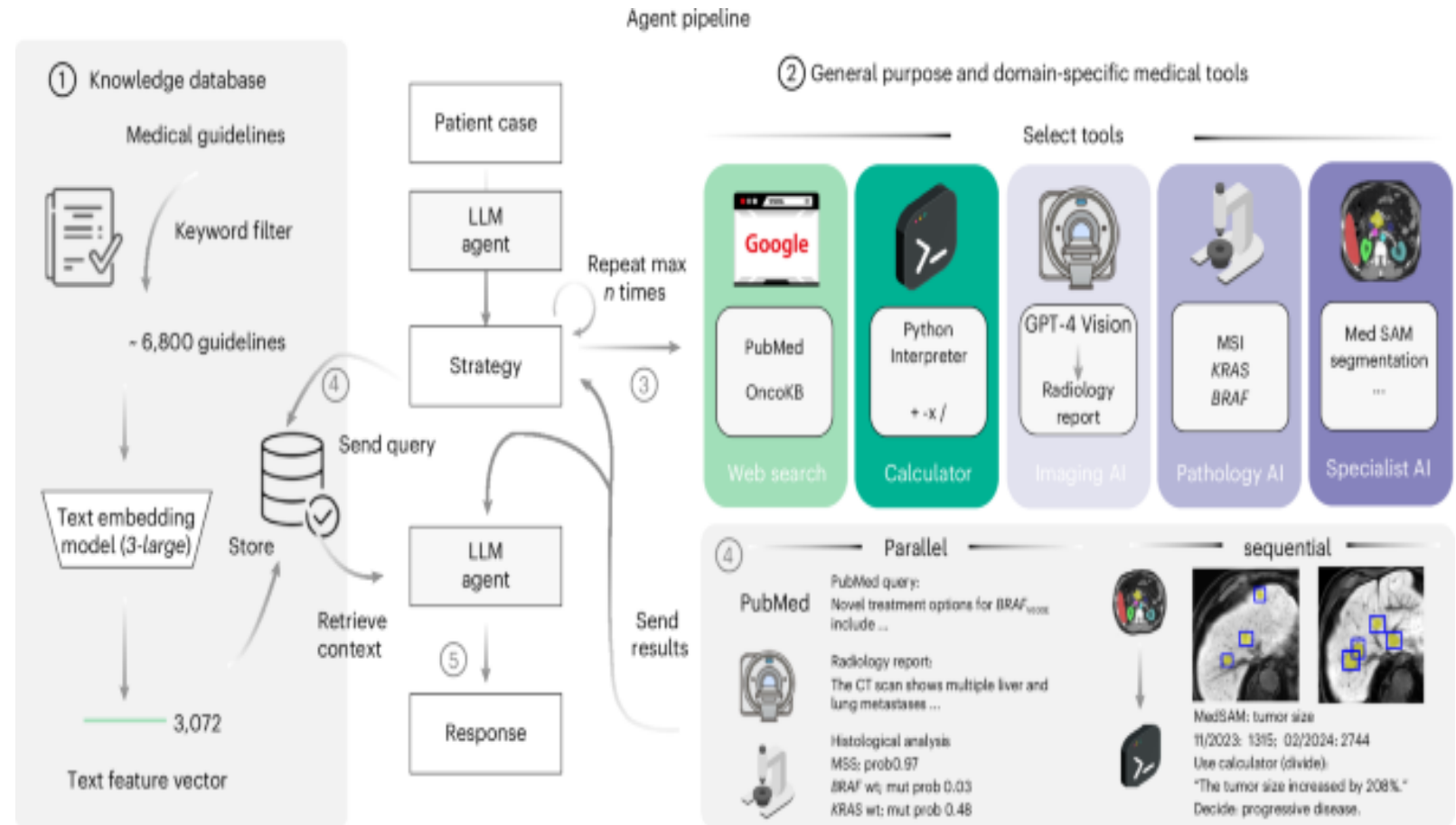
Best first deployment model: bounded tasks, traceable evidence, human oversight, and no illusion of autonomous medical judgment.



Tumour-board agents and multimodal decision support

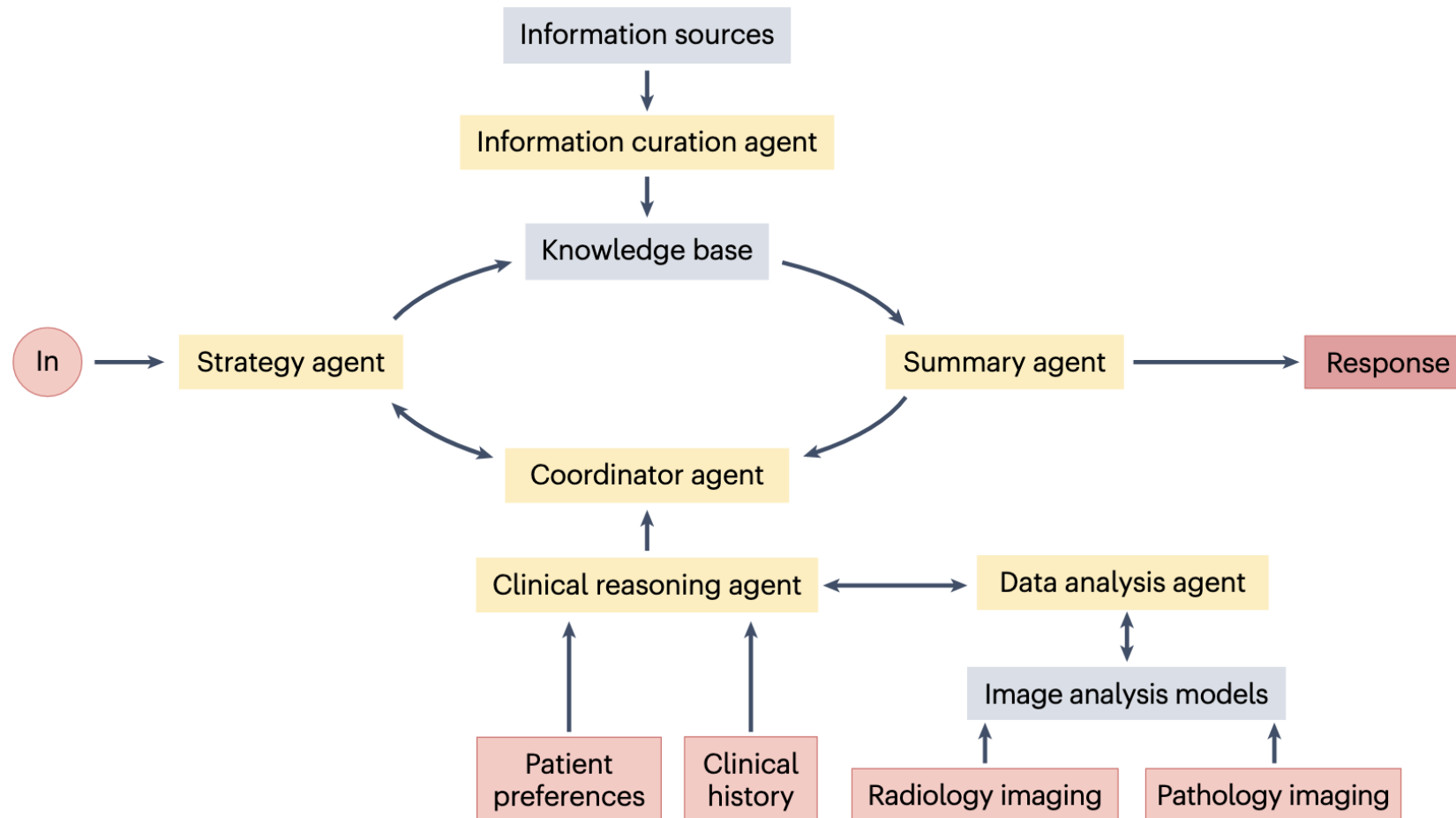
What the Nature Cancer 2025 study showed

A multimodal oncology AI agent used pathology, radiology, web search tools, and precision-oncology resources to support decision-making on 20 realistic cases.



Ferber et al., Nature Cancer 2025

Multi-agent framework for oncological treatment decisions



Truhn et al., Nature Reviews Cancer 2026

Trial matching: one of the most tangible near-term use cases

Why this is hard

- Eligibility criteria are buried in long protocols
- Relevant information lives in free text and unstructured reports
- Biomarkers, prior lines, washout periods, and sequencing rules interact
- The decision window is often short

Patient
chart
+ genomics



Eligibility
abstraction
and reasoning



Ranked trial
options

What a good matching tool should do

Not merely keyword search. It should retrieve relevant trials, reason over structured and unstructured eligibility criteria, explain why a patient is or is not a fit, and preserve human oversight.

TrialGPT: a concrete example of LLM-assisted patient-to-trial matching

Retrieval

Matching

Ranking

>90%

relevant trials recalled using
<6% of the initial collection

87.3%

criterion-level matching accuracy in manual evaluation

-42.6%

screening time in the user study

Why this paper is useful for teaching

It breaks the task into clinically intuitive steps: find candidate trials, reason over each criterion, then generate a trial-level ranking.

But remember the realism check

Benchmark success is not the same as clinical deployment. Workflow integration, updated trial databases, chart quality, and supervision still determine real-world value.

Digital model vs digital shadow vs digital twin

Digital model

Standalone simulation
No ongoing connection to the patient

Physical -> Digital: ✗

Digital -> Physical: ✗

Dynamic updating: Usually
no

Digital shadow

Patient updates the model
But the model does not feed back

Physical -> Digital: ✓

Digital -> Physical: ✗

Dynamic updating:
Monitoring

Digital twin

Bidirectional, dynamic,
predictive
Used to inform decisions

Physical -> Digital: ✓

Digital -> Physical: ✓

Dynamic updating: Yes

This distinction matters because much of the current literature uses “digital twin” more loosely than the strict definition allows.

What a cancer patient digital twin is trying to become

iplinary
Cancers

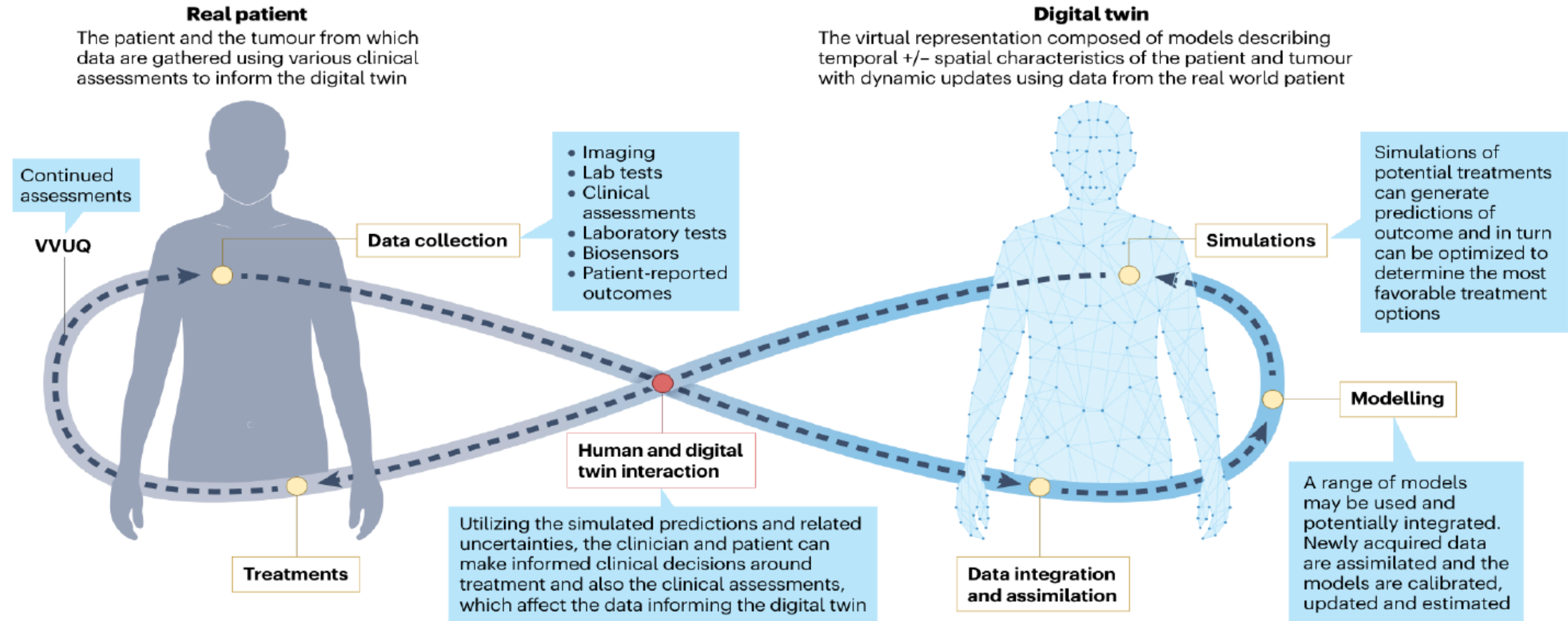


Fig. 1 | Example of a digital twin of a patient with cancer and the tumour. +/-, and/or; VVUQ, verification, validation and uncertainty quantification. Adapted with

Ashgar et al., Nature Reviews Cancer 2025

**Most current systems
are not here yet**

Synthetic data: useful, but not the same thing as a digital twin

What synthetic data are

Artificially generated data designed to mimic the statistical structure of real cohorts while reducing direct disclosure of real patient records.

Main uses

Data sharing
Benchmarking
Augmentation
Method development

How they should be evaluated

Fidelity: do they preserve the right distributions and relations?
Utility: do they behave correctly in downstream tasks?
Privacy: do they avoid re-identification risk?

Synthetic data are cohort-level surrogates. A digital twin is a patient-level dynamic model. These are not interchangeable concepts.

Synthetic control arms: evidence from oncology trials

Growing proof-of-concept evidence that AI-generated controls can replicate real trial outcomes.

AML — Eckardt et al. 2024 (npj Digit Med)

Synthetic data from 4 pooled AML trials closely replicate real disease biology and outcomes using CTAB-GAN+ and Normalizing Flows. Prognostic variables overlap with real IPSS-M discovery cohort.

Colorectal — Azizi et al. 2021 (BMJ Open)

Synthetic controls replicate key findings of the phase III N0147 trial (FOLFOX ± cetuximab). High correspondence in hazard ratios and significance levels vs. original secondary analysis.

NSCLC — Elvatun et al. 2025 (PLoS Digit Health)

6 generative models compared. PrivBayes and CTGAN closely match real survival curves and treatment effects. Performance degrades after ~25 months — a known limitation of sparse late data.

Breast — El-Kababji et al. 2025 (JMIR)

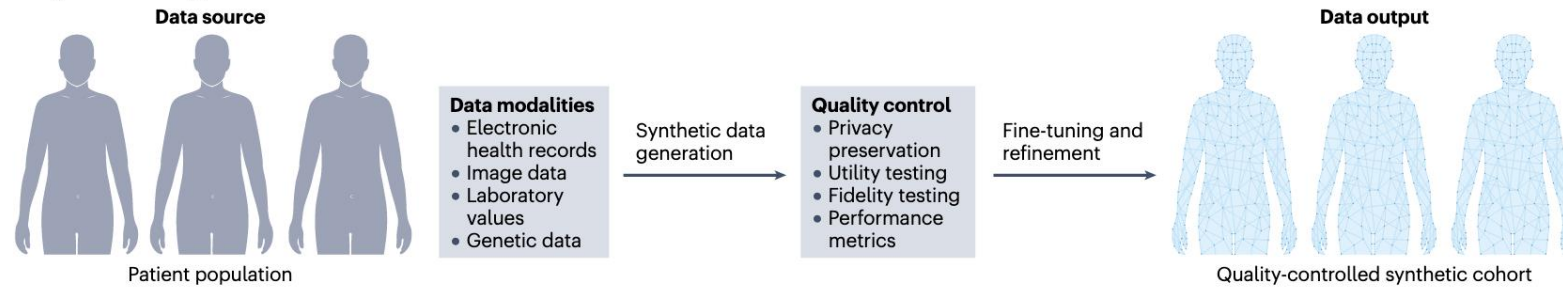
9 fully recruited breast cancer trials. Synthetic patients trained on early enrollees reliably reproduce full datasets, mitigating slow accrual and underpowered analyses.

Fidelity, utility, and privacy must all be assessed. Synthetic controls should ideally be generated by an independent party to prevent conflicts of interest (Eckardt et al. 2026).

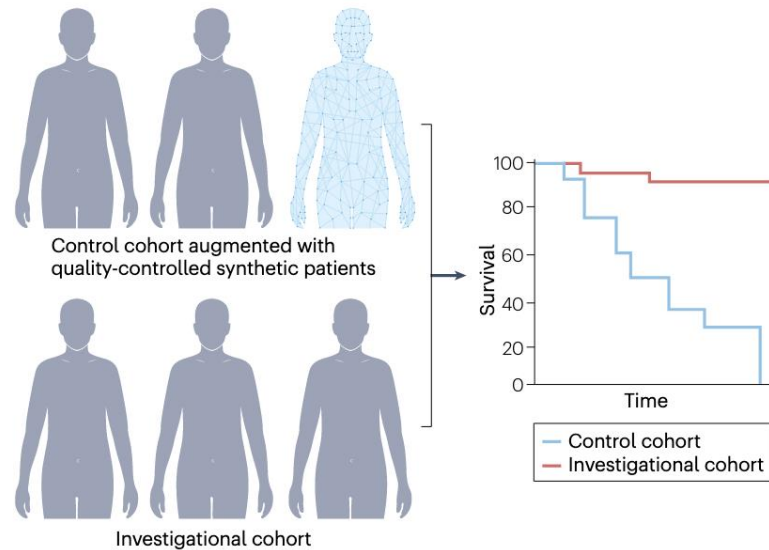
Synthetic control cohorts: promise and pitfalls for trials

13th Belgian Multidisciplinary Meeting on Urological Cancers

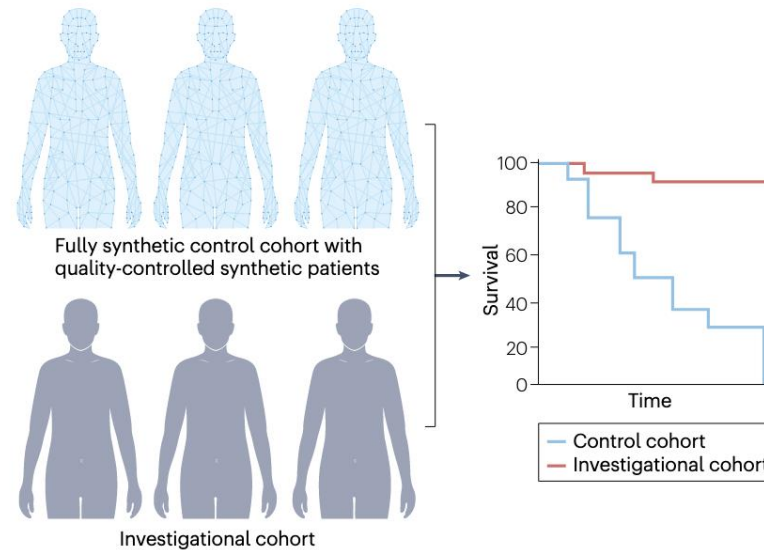
a Synthetic data generation



b Control cohort augmentation



c Control cohort substitution



Potential value: rare tumours, recruitment

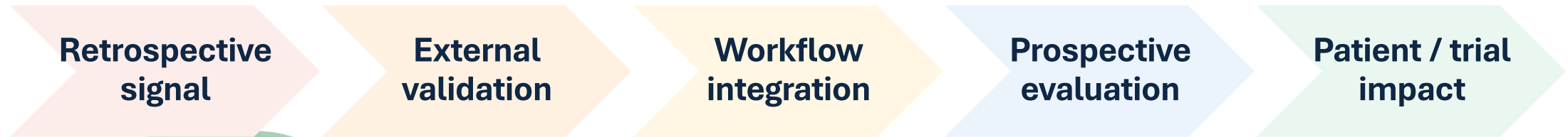
- constraints, reduced control-arm burden, method development.

Major risks: confounding, dataset shift, temporal drift, hidden bias, and overconfidence in “replacement” of randomized evidence.

Most realistic role today: supporting evidence

- generation, augmentation, and simulation - not magic substitution for rigorous trial design.

From hype to evidence: crossing the “AI chasm”



Why many models fail to cross it

Single-center data
Poor transportability
Silent dataset shift
Messy real-world inputs

What prospective reality adds

User behavior
Automation bias
Time pressure
Interoperability constraints
Governance requirements

What future winners will have in common

External validation
Traceable evidence
Human oversight
Clinical workflow fit
Measured benefit

We should be optimistic — but disciplined. In oncology, the real question is not “Can AI do this?” but “Has it been validated where it matters?”

Take-home messages

1

AI is shifting from isolated prediction models toward an infrastructure layer for cancer research.

2

The most immediate therapeutic impact is upstream: drug development, dose optimization, and trial operations.

3

GU oncology is especially AI-ready: multimodal, longitudinal, biomarker-rich data across RCC, prostate, and bladder.

4

Agents will probably arrive in bounded workflows before true patient digital twins arrive in routine practice.

5

Validation matters most.

Inspirational but realistic conclusion: expect better copilots, smarter matching, more adaptive trial systems, and richer model-informed development well before fully autonomous oncology.



SAVE THE DATE

AI in
ONCOLOGY

By IHU PRISM

22-23 June, 2026

Novotel Tour Eiffel - Paris

2nd Edition



Acknowledgments



Fabrice Barlesi
Carine Ngo
Axel Le Cesne
Benjamin Verret

Fabrice André
Maud Kamal
Hakim Benkirane
Paul-Henry
Cournède

Christophe Massard
Yohann Loriot
Stefan Fröhling
Jakob Kather



Thank you for your attention!

