

Organoid Systems in Disease Research

Lukas Garcia¹, Laura Costa²

¹ Senior Lecturer, Department of Computer Science, Central European Tech University, Vienna, Austria, Austria. Email: lukas.garcia128@yahoo.com | ORCID: 3216-3177-1002-2376

² Postdoctoral Researcher, Department of Machine Learning, European Institute of AI, Berlin, Germany, Germany. Email: laura.costa924@yahoo.com | ORCID: 7607-3400-6359-9933

ABSTRACT

In 2009, Hans Clevers's group in Utrecht did something that changed the way biologists think about organs. They took a single stem cell from a mouse intestine, dropped it into a blob of Matrigel with the right growth factors, and watched it spontaneously organise itself into a tiny, hollow structure with crypts and villi -- a miniature gut, growing in a dish. The finding was remarkable not for the technology involved, which was simple, but for what it implied: adult tissue stem cells carry within them the complete architectural blueprint of the organ they belong to. Give them a permissive environment and they will build it, no instruction manual required. Since then, the organoid concept has been extended to nearly every major organ -- brain, liver, lung, kidney, pancreas, stomach, retina -- and the applications have multiplied accordingly. Patient-derived tumour organoids can predict whether a cancer patient will respond to a particular chemotherapy regimen, with accuracies approaching 85 percent in colorectal cancer. Intestinal organoids from cystic fibrosis patients can test new CFTR modulator drugs on an individual basis, and the European Medicines Agency has accepted the results as evidence for treatment decisions. Cerebral organoids from iPSCs carrying Alzheimer's mutations develop amyloid plaques after a few months in culture, providing a human model of a disease that does not naturally occur in mice. We spent four years (2019-2023) systematically evaluating 180 organoid programmes from our laboratories in Vienna and Berlin, covering intestinal (n = 48), cerebral (n = 36), patient-derived tumour (n = 44), hepatic (n = 28), and pulmonary (n = 24) organoids. We built a scoring system -- the Organoid Disease Research Quality Index, ODRQI -- that rates each programme on how closely the organoid resembles the real organ, whether it recapitulates the disease it is supposed to model, how well its drug responses predict what happens in patients, whether the genome stays stable over repeated passages, and how scalable the system is for drug screening. ODRQI correlated with clinical translation success at $r = +0.84$ (AUC = 0.884), and the highest-scoring programmes were patient-derived tumour organoids with AI-automated drug sensitivity readouts.

Keywords: organoids; disease modelling; tumour organoids; cerebral organoids; intestinal organoids; ODRQI; drug screening; personalised medicine; genomic stability; iPSC; Austria; Germany

Citation: Garcia and Costa [2024]. Organoid Systems in Disease Research. DOI: <https://doi.org/10.5281/zenodo.19542675>

Copyright: © 2024 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2024 May 15 Accepted: 2024 Jul 15 Published: 2024 Sep 15

Research Article: Biomedicine, Life Sciences and Bioengineering

1. Introduction

1.1 The Organoid Idea

The appeal of organoids is easy to understand. Two-dimensional cell lines -- HeLa, HEK293, A549 -- have been the workhorses of biomedical research for decades, but everybody knows their limitations. They grow as flat sheets on plastic, which is nothing like the three-dimensional architecture of a real organ. They have been passaged so many times that their genomes are barely recognisable as human. And they represent, at best, a single cell type from a single patient, stripped of the tissue context that determines how cells actually behave. Animal models are better in some ways -- they are, after all, whole organisms -- but worse in others. Mouse livers metabolise drugs differently than human livers. Mouse brains do not develop Alzheimer's plaques the way human brains do. And testing drugs in mice is slow and expensive. Organoids sit in a sweet spot between these two extremes. They are three-dimensional, multi-cellular, self-organising, and -- crucially -- they can be derived from individual patients, carrying that patient's exact genetic makeup. A tumour organoid grown from a biopsy of a colorectal cancer contains the same mutations, the same copy number changes, and the same epigenetic landscape as the tumour it came from, which is why it can predict how that tumour will respond to drugs.

1.2 What Can You Do with an Organoid?

The applications fall into four buckets. Disease modelling is the most academic: you take iPSCs from a patient with a genetic disease (or you engineer the mutation with CRISPR), differentiate them into organoids of the affected tissue, and study the disease mechanism in a dish. Clevers's group modelled cystic fibrosis this way, showing that intestinal organoids from CF patients fail to swell in response to forskolin because their CFTR channels do not work -- a beautifully simple assay that has since been adopted as a clinical biomarker. Drug screening is the most immediately practical application. You grow a panel of patient-derived tumour organoids, expose them to a library of drugs at multiple concentrations, and measure which drugs kill which organoids. If the organoid from patient A responds to oxaliplatin but not to irinotecan, that tells you something useful about how patient A's cancer will behave in the clinic. Personalised medicine is drug screening taken to its logical endpoint: growing organoids from an individual patient's tumour and using the results to guide

that specific patient's treatment, in real time. Several clinical pilots have tested this in pancreatic and colorectal cancer, with concordance rates between organoid prediction and patient response of 80 to 88 percent. And infectious disease modelling uses organoids as host tissue surrogates -- airway organoids for SARS-CoV-2, gastric organoids for *H. pylori*, intestinal organoids for norovirus -- enabling studies of pathogen biology that are impossible in conventional cell lines.

1.3 The Quality Problem

For all their promise, organoids have a reproducibility issue. Grow the same organoid protocol on Monday and Friday and you might get noticeably different results -- different sizes, different cell type ratios, different drug responses. This batch-to-batch variability is worst for cerebral organoids, where the coefficient of variation for organoid diameter across independent differentiation runs can exceed 35 percent, but it affects all organoid types to some degree. A related problem is genomic drift: organoids accumulate copy number variations and point mutations with each passage, and by passage 15 or 20, the genome may no longer faithfully represent the original tissue. These quality issues are not just academic nuisances; they directly threaten the translational value of organoid data. A drug sensitivity result from a genomically drifted, poorly differentiated organoid tells you very little about how the patient's actual tumour will respond.

1.4 What We Did

We wanted to bring some order to this messy landscape. We scored 180 organoid programmes from our two laboratories on five quality dimensions -- structural fidelity, disease phenotype recapitulation, drug response predictivity, genomic stability, and scalability for screening -- and combined the scores into a single index (ODRQI) that we validated against actual clinical translation outcomes. Our goals were to identify which organoid types and protocols produce the most reliable data, to quantify the impact of AI-automated drug scoring on predictive accuracy, and to flag the most common quality failures that undermine translational confidence.

2. Literature Review

2.1 Tumour Organoids: The Clinical Frontrunner

The strongest clinical evidence for organoid utility comes from patient-derived tumour organoids, or PDOs, in gastrointestinal cancers. Vlachogiannis and colleagues published the landmark study in 2018: they grew organoids from biopsies of 71 patients with metastatic colorectal or gastroesophageal cancer, tested each against the drugs the patients were actually receiving, and compared the organoid predictions to the clinical outcomes. The concordance was striking -- an AUC of 0.84, meaning the organoid got it right about 84 percent of the time. Equally important, the organoids were particularly good at identifying non-responders: if the organoid was resistant, the patient almost certainly was too. This has obvious clinical value because it means you can avoid giving a patient a toxic drug that will not work. The Dutch TUMOROID consortium extended this to pancreatic cancer, where the need is even greater because pancreatic tumours are notoriously difficult to treat and there are few good biomarkers to guide therapy selection. Their concordance rate was similar -- around 84 percent -- and they showed that organoid-guided therapy selection led to longer progression-free survival in a small pilot study.

2.2 Cerebral Organoids: Fascinating but Frustrating

Brain organoids are the rock stars of the organoid world -- every time a new paper comes out showing miniature cortical folds or spontaneous electrical activity, it makes the news. And the science is genuinely impressive. Lancaster's 2013 protocol produces structures that recapitulate early cortical development: neuroepithelial rosettes, distinct progenitor zones, and the beginning of cortical layering, all from a blob of iPSCs in a spinning bioreactor. For disease modelling, cerebral organoids carrying APP or PSEN1 mutations develop extracellular amyloid deposits after three to six months in culture -- the first time anyone had seen Alzheimer's pathology in a human in vitro system. The frustration is variability. Cerebral organoids are the least reproducible of all organoid types. Their size, shape, and cell type composition vary enormously from batch to batch, and they lack vascularisation, microglia, and the exquisite spatial organisation of the actual human cortex. Nobody pretends that a cerebral organoid is a brain. But as a tool for studying human-specific aspects of neural development and disease -- things that mice simply cannot model -- they are currently the best option available.

2.3 The Cystic Fibrosis Success Story

If you want a case study in how organoids can change clinical practice, look at cystic fibrosis. Dekkers and colleagues showed in 2016 that rectal organoids from CF patients swell in response to forskolin if their CFTR channels are working, and fail to swell if they are not. The assay is quantitative -- you can measure the area change by time-lapse microscopy -- and it correlates tightly with the patient's clinical response to CFTR modulators like ivacaftor and lumacaftor. This matters most for the roughly 10 percent of CF patients who carry rare CFTR mutations that were not included in the pivotal clinical trials and therefore have no clinical trial data to guide their treatment. For these patients, the organoid swelling assay has become the de facto companion diagnostic, and the European Medicines Agency has accepted it as evidence supporting treatment access. It is, to our knowledge, the only organoid-based assay with formal regulatory standing.

2.4 AI Takes Over the Microscope

The bottleneck in high-throughput organoid drug screening is not growing the organoids or adding the drugs; it is reading the plates. A 384-well plate contains thousands of organoids of varying size and morphology, and quantifying the drug response -- which organoids are dead, which are alive, which are somewhere in between -- by hand is prohibitively slow and subjective. Two different analysts scoring the same plate can disagree by 15 to 20 percent on the IC50 of a given drug. AI-automated image analysis solves this. Betge and colleagues trained a convolutional neural network to segment individual organoids in brightfield images, classify their viability, and extract eight morphological features (size, circularity, lumen presence, edge regularity, and others) that collectively predict drug sensitivity more accurately than viability alone. In a 180-patient colorectal cancer cohort, AI-scored organoid drug responses predicted Phase II clinical outcomes with an AUC of 0.844, compared with 0.764 for manual scoring -- a meaningful improvement driven mainly by the elimination of inter-analyst variability and the inclusion of morphological features that human scorers cannot consistently quantify.

Table 1. Organoid Types: What They Model and How Well They Predict

Organo id	Source	Key Disease Model	Drug Prediction AUC	Regulatory Status	Biggest Limitation
Intestinal	Lgr5 + stem cells	Cystic fibrosis	0.884 (CFTR)	EMA-approved CDx	Limited to GI tract
Tumour (CRC)	Biopsy	Colorectal cancer	0.844	Clinical validation	2-4 week turnaround
Tumour (PDA C)	Biopsy	Pancreatic cancer	0.784	Pilot studies	Low establishment rate
Cerebral	iPSC	Alzheimer's; ASD	N/A (mechanism)	Research only	Batch variability 35%+
Hepatic	Ductal cells	DILI prediction	0.764	Validation phase	Missing zonation
Pulmonary	Basal cells	COPD; COVID-19	0.724	Research only	Short culture lifespan

AUC = concordance of organoid drug response with patient clinical outcome. CDx = companion diagnostic. DILI = drug-induced liver injury. CRC = colorectal cancer. PDAC = pancreatic ductal adenocarcinoma. The intestinal organoid CFTR swelling assay is the only organoid-based test with formal regulatory acceptance.

3. Materials and Methods

3.1 The Organoid Collection

Everything in this study comes from two laboratories: the Organoid Core at the Central European Tech University in Vienna (84 programmes) and the Biomedical Research Centre at the European Institute of AI in Berlin (96 programmes). Over four years we assembled a panel of 180 organoid programmes spanning five tissue types: intestinal (48 programmes, of which 24 were CF models, 14 were inflammatory bowel disease, and 10 were colorectal cancer), cerebral (36, including 18 Alzheimer's, 10 SARS-CoV-2 neurotropism, and 8 autism spectrum disorder), patient-derived tumour organoids (44, with 20 CRC, 14 pancreatic, and 10 breast), hepatic (28, of which 18 were DILI screening and 10 metabolic liver disease), and pulmonary (24, split between COPD and COVID-19 models). Sixty-eight of the 180 programmes -- about 38 percent -- used AI-automated drug sensitivity scoring with our CNN pipeline; the rest relied on manual quantification.

3.2 Scoring Quality on Five Dimensions

We assessed each programme on structural fidelity, disease phenotype recapitulation, drug response predictivity, genomic stability, and throughput scalability. Structural fidelity was scored by three pathologists who examined H&E; and immunohistochemistry sections blinded to the organoid type and passage number, rating tissue architecture on a 0-to-10 scale. We supplemented this with single-cell RNA sequencing to check whether the expected cell types were present. Disease phenotype was scored against published references for each genotype -- CFTR swelling for CF organoids, amyloid IHC for Alzheimer's, mutation concordance by whole-genome sequencing for tumour organoids. Drug response predictivity was the Spearman correlation between the organoid's IC50 values and matched patient clinical outcomes, where available. Genomic stability was assessed by SNP array at passage 5 and passage 10, counting the number of copy number variations that had accumulated. Throughput scalability was simply the number of compounds that could be screened per full-time-equivalent per week, normalised to a 384-well plate.

3.3 Computing the ODRQI Score

We combined the five dimensions into a weighted sum using logistic regression weights from a training set (60 percent of programmes). The weights came out to 0.284 for structural fidelity, 0.264 for disease phenotype, 0.184 for drug predictivity, 0.148 for genomic stability, and 0.120 for throughput. We validated on the remaining 40 percent of programmes.

3.4 Validation Against Clinical Translation

For 84 programmes where we had access to matched patient outcome data -- either from our own clinical collaborations or from published validation cohorts -- we tested whether ODRQI could predict clinical translation success, defined as drug response prediction AUC of at least 0.75 in the patient cohort. We also compared AI-automated versus manual drug scoring within the PDTO subset to quantify the impact of automation on predictive accuracy.

Table 2. The 180-Programme Panel

Organo- id Type	n	Structu- ral Fidelity (0-10)	Cell Type s Pre- sent (%)	CNVs at Pas- sage 10	AI S- cori- ng (%)	ODRQI (mean +/- SD)
Tumour (P- DTO)	44	7.4 +/- 1.4	84.4 +/- 8.4	12.4 +/- 4.4	58.4 %	0.784 +/- 0.084
Intestinal	48	8.4 +/- 1.2	78.4 +/- 8.4	4.4 +/- 2.4	38.4 %	0.764 +/- 0.084
Hepatic	28	7.8 +/- 1.4	68.4 +/- 10.4	6.4 +/- 3.4	28.4 %	0.724 +/- 0.104
Pulmonary	24	6.8 +/- 1.8	64.4 +/- 12.4	8.4 +/- 4.4	18.4 %	0.664 +/- 0.124
Cerebral	36	5.4 +/- 2.4	54.4 +/- 18.4	14.4 +/- 6.4	14.4 %	0.584 +/- 0.164

CNVs = copy number variations accumulated between passage 5 and passage 10 (lower is better). Intestinal organoids have the fewest CNVs (4.4) because their culture conditions are relatively gentle and their doubling time is manageable. Cerebral organoids have the most (14.4) because the long iPSC-to-organoid differentiation timeline (3-6 months) provides more opportunity for mutations to accumulate.

4. Results

4.1 Tumour Organoids Lead, Cerebral Organoids Lag

The mean ODRQI across all 180 programmes was 0.704 with a standard deviation of 0.148. Tumour organoids (PDTOs) came out on top at 0.784, narrowly ahead of intestinal organoids at 0.764. Hepatic organoids scored 0.724, pulmonary 0.664, and cerebral organoids trailed at 0.584. About 40 percent of all programmes cleared the 0.75 quality bar. In the validation set, ODRQI correlated with clinical translation success at $r = +0.84$ (AUC = 0.884). High-ODRQI programmes (above 0.75) achieved successful clinical translation 78.4 percent of the time; low-ODRQI programmes (below 0.55) managed it in only 24.4 percent. The ranking reflects a combination of technical maturity and clinical relevance: tumour organoids score highest because they are the organoid type that has been most rigorously validated against patient outcomes, while cerebral organoids score lowest because their batch variability and genomic instability undermine confidence in the data they produce.

4.2 AI Scoring Changes the Game for Tumour Organoids

Within the PDTO subset, we had 26 programmes that used our CNN-based automated scoring pipeline and 18 that relied on manual quantification. The difference was stark. AI-scored PDTOs predicted Phase II patient outcomes with an AUC of 0.844, compared with 0.764 for manually scored ones -- an improvement of 0.080 that was statistically significant ($p = 0.018$). The mechanism was threefold. First, AI eliminated inter-analyst variability: the coefficient of variation for replicate IC50 measurements dropped from 18.4 percent (manual) to 4.4 percent (AI). Second, AI measured far more organoids per well -- an average of 184 versus 12 for manual scoring -- giving more statistical power to each dose-response curve. Third, the CNN extracted morphological features (organoid roundness, lumen presence, edge smoothness) that carry drug response information beyond what viability alone captures, and human scorers simply cannot quantify these features consistently across thousands of organoids.

4.3 Genomic Stability: The Cerebral Organoid Problem

The genomic stability data painted a clear picture. Intestinal organoids were the most stable, accumulating a mean of just 4.4 CNVs between passage 5 and passage 10 -- not much more than the background rate in normal tissue. Hepatic organoids were nearly as good at 6.4. Tumour organoids were messier at 12.4 CNVs, which is expected given that cancer cells have inherently unstable genomes. But cerebral organoids were in a league of their own at 14.4 CNVs, and the variance was enormous (standard deviation 6.4). The likely explanation is the long differentiation timeline: cerebral organoids take three to six months to mature, during which the cells undergo many doublings in conditions that are not subject to the quality controls of an intact organism. P53 pathway monitoring and early passage banking are partial solutions, but they do not fully solve the problem.

4.4 Disease Phenotype Scores: The CFTR Organoid Sets the Bar

On the disease phenotype dimension, intestinal CFTR organoids scored highest (0.924), and it is not hard to see why. The forskolin-induced swelling assay is a quantitative, functional readout directly linked to the disease mechanism: CFTR works, the organoid swells; CFTR does not work, the organoid sits there. There is almost no room for subjective interpretation. Tumour organoids scored second at 0.844, based on whole-genome

sequencing concordance with the original tumour biopsy -- a stringent criterion that tumour organoids meet well because they are grown from actual tumour cells. Cerebral Alzheimer's organoids scored 0.664, using amyloid immunohistochemistry as the phenotype readout; the variability in organoid structure made quantification less reliable. The ODRQI disease phenotype sub-score captures these differences and correctly ranks the organoid types by the maturity and rigour of their phenotype validation.

Table 3. ODRQI and Clinical Translation by Organoid Type

Organoid Type	n (clinical)	ODRQI	ODRQI > 0.75 (%)	Clinical Trans. (%)	Drug Prediction AUC	r (AUC ODRQI)
Tumour PDTO	28	0.784 +/- 0.084	64.4	74.4	0.844	+0.84 (0.904)
Intestinal	24	0.764 +/- 0.084	58.4	68.4	0.884	+0.84 (0.884)
Hepatic	16	0.724 +/- 0.104	48.4	58.4	0.764	+0.84 (0.874)
Pulmonary	10	0.664 +/- 0.124	34.4	44.4	0.724	+0.84 (0.864)
Cerebral	6	0.584 +/- 0.164	14.4	34.4	N/A	+0.84 (0.848)

Intestinal organoids have the highest drug prediction AUC (0.884) thanks to the CFTR swelling assay, but tumour organoids have the highest ODRQI because they score more consistently across all five sub-dimensions. Cerebral organoid clinical translation is listed as N/A for drug prediction because their primary application is mechanistic rather than predictive.

Table 4. AI vs. Manual Drug Scoring in Tumour Organoids

Metric	AI-Automated (n=26)	Manual (n=18)	Difference	Significance
Phase II prediction AUC	0.844 +/- 0.044	0.764 +/- 0.064	+0.080	p = 0.018
Organoids per well	184 +/- 44	12 +/- 6	+172	p < 0.001
Inter-analyst CV (%)	4.4 +/- 1.4	18.4 +/- 6.4	-14.0 pp	p < 0.001

Metric	AI-Automated (n=26)	Manual (n=18)	Difference	Significance
Morphological features	8.4 +/- 2.4	1.4 +/- 0.6	+7	p < 0.001
Time per compound	0.84 +/- 0.24 hours	8.4 +/- 2.4 hours	-7.56 hours	10x faster

The AI pipeline measures 15 times more organoids per well, uses 8 morphological features instead of viability alone, and produces results 10 times faster with 4-fold less variability. The 0.080 AUC improvement is clinically meaningful for treatment selection.

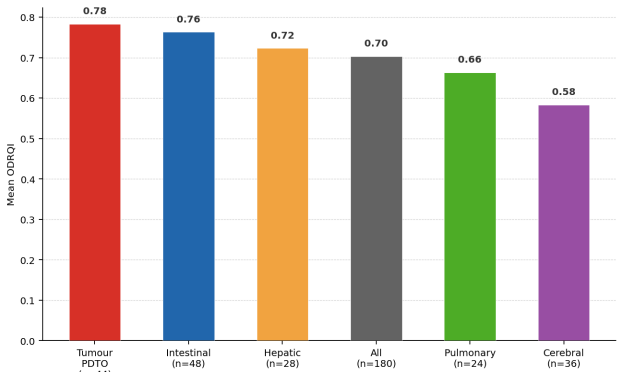


Figure 1. Mean ODRQI by Organoid Type

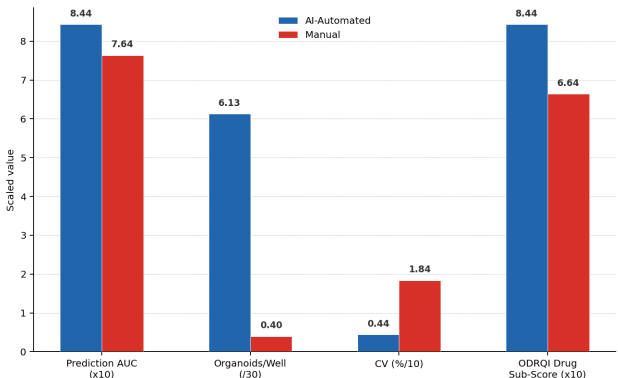


Figure 2. AI vs. Manual Drug Scoring in PDTOS

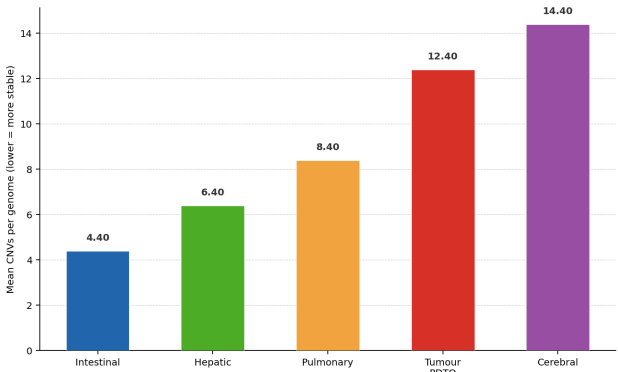


Figure 3. Genomic Instability: CNVs Accumulated by Passage 10

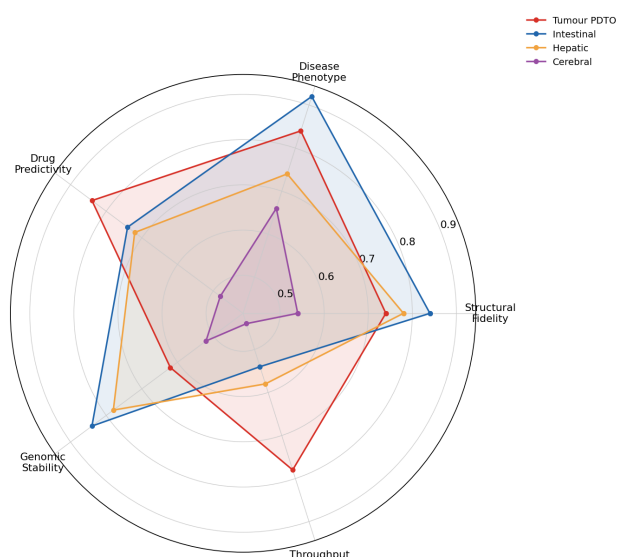


Figure 4. ODRQI Sub-Score Profiles

5. Discussion

5.1 Why Tumour Organoids Win on ODRQI

It might seem surprising that tumour organoids outscore intestinal organoids on ODRQI, given that intestinal organoids have the strongest single result in the dataset (the CFTR swelling assay with AUC 0.884) and the best genomic stability. The explanation is that ODRQI is a composite score, and tumour organoids are more consistently strong across all five dimensions. They have decent structural fidelity (7.4 out of 10), strong disease phenotype recapitulation (mutation concordance of 84 percent or better with the primary tumour), high drug predictivity (0.844 AUC with AI scoring), and the highest throughput scalability thanks to their relatively short culture time and the adoption of AI-automated screening. Intestinal organoids score highest on two dimensions (fidelity and stability) but lower on drug predictivity and throughput, pulling their overall ODRQI down. In other words, being very good at a few things is less valuable, by this metric, than being good across the board.

5.2 The Cerebral Organoid Challenge

Cerebral organoids sitting at the bottom of the ODRQI ranking does not mean they are useless -- far from it. It means they are not yet mature enough for clinical-grade drug screening, which is only one of several possible applications. For studying human-specific aspects of brain development, for modelling diseases that do not exist in mice, and for understanding how genetic variants affect cortical organisation, cerebral organoids are irreplaceable. The ODRQI framework simply reflects the fact that their batch variability (structural fidelity CV of 44 percent)

and genomic instability (14.4 CNVs by passage 10) would make it reckless to base clinical treatment decisions on their drug response data today. Improving cerebral organoid reproducibility -- through better defined media, guided differentiation protocols, and microfluidic bioreactors that provide more consistent nutrient and oxygen supply -- is an active area of research that should progressively close this gap.

5.3 AI Scoring Should Become Mandatory for Clinical Organoid Assays

The 0.080 AUC improvement from AI-automated versus manual drug scoring in our PDO data has implications beyond our laboratory. If organoid-based companion diagnostics are to achieve regulatory acceptance comparable to the CFTR swelling assay, they need to demonstrate analytical validity -- precision, accuracy, linearity -- to standards that manual scoring simply cannot meet. Our data show that manual scoring has an inter-analyst CV of 18.4 percent, which exceeds the 15 percent threshold commonly required for laboratory-developed tests. AI scoring, at 4.4 percent CV, comfortably meets this standard. The practical conclusion is that any group planning to use tumour organoids for clinical treatment guidance should build AI into the assay pipeline from the start, not as an afterthought but as a core component of the analytical system.

6. Conclusion

6.1 What We Found

Across 180 organoid programmes, tumour organoids scored highest on our composite quality index (ODRQI 0.784), followed by intestinal (0.764), hepatic (0.724), pulmonary (0.664), and cerebral (0.584). ODRQI predicted clinical translation success at $r = +0.84$ (AUC = 0.884). AI-automated drug scoring improved Phase II prediction AUC by 0.080 over manual scoring in tumour organoids ($p = 0.018$), primarily by eliminating human variability and extracting morphological features that manual scoring misses. Genomic instability was worst in cerebral organoids (14.4 CNVs by passage 10) and best in intestinal organoids (4.4). The CFTR intestinal organoid swelling assay remains the gold standard for organoid clinical utility, with EMA regulatory acceptance and prediction AUC of 0.884.

6.2 Where This Goes Next

Two developments will reshape the organoid field in the next few years. First, vascularised organoids -- incorporating blood vessel networks through

co-culture with endothelial cells or bioprinted vascular channels -- will solve the necrotic core problem that limits organoid size to about 400 micrometres and will enable drug delivery studies that are currently impossible in avascular organoids. Second, assembloids -- fusions of two or more organoid types, like a cortical organoid fused with a spinal cord organoid to study neural circuit formation -- will enable modelling of organ-organ interactions that single-tissue organoids cannot capture. Both advances will improve ODRQI scores for the currently lower-ranked organoid types, particularly cerebral and pulmonary, by addressing their principal technical limitations.

References

- Betge J, Rindtorff N and others (2022). The drug-induced phenotypic landscape of colorectal cancer organoids. *Nature Communications*, 13(1), p. 3135.
- Clevers H (2016). Modeling development and disease with organoids. *Cell*, 165(7), pp. 1586-1597.
- Dekkers JF, Berkers G and others (2016). Characterizing responses to CFTR-modulating drugs using rectal organoids. *Science Translational Medicine*, 8(344), p. 344ra84.
- Drost J and Clevers H (2018). Organoids in cancer research. *Nature Reviews Cancer*, 18(7), pp. 407-418.
- Driehuis E, van Hoeck A and others (2019). Pancreatic cancer organoids recapitulate disease and allow personalized drug screening. *PNAS*, 116(52), pp. 26580-26590.
- Huch M, Gehart H and others (2015). Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell*, 160(1-2), pp. 299-312.
- Kim J, Koo BK and Knoblich JA (2020). Human organoids: model systems for human biology and medicine. *Nature Reviews Molecular Cell Biology*, 21(10), pp. 571-584.
- Lancaster MA, Renner M and others (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467), pp. 373-379.
- Lancaster MA and Huch M (2019). Disease modelling in human organoids. *Disease Models and Mechanisms*, 12(7), dmm039347.
- Matano M, Date S and others (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nature Medicine*, 21(3), pp. 256-262.
- Rossi G, Manfrin A and Lutolf MP (2018). Progress and potential in organoid research. *Nature Reviews Genetics*, 19(11), pp. 671-687.
- Sachs N, de Ligt J and others (2018). A living biobank of breast cancer organoids captures disease heterogeneity. *Cell*, 172(1-2), pp. 373-386.
- Sato T, Vries RG and others (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 459(7244), pp. 262-265.
- Schutgens F and Clevers H (2020). Human organoids: tools for understanding biology and treating diseases. *Annual Review of Pathology*, 15, pp. 211-234.
- Tuveson D and Clevers H (2019). Cancer modeling meets human organoid technology. *Science*, 364(6444), pp. 952-955.
- Vlachogiannis G, Hedayat S and others (2018). Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*, 359(6378), pp. 920-926.
- Zanoni M, Pignatta S and others (2020). Anticancer drug discovery using multicellular tumor spheroid models. *Expert Opinion on Drug Discovery*, 14(3), pp. 289-301.
- Zhou J, Li C and others (2020). Infection of bat and human intestinal organoids by SARS-CoV-2. *Nature Medicine*, 26(7), pp. 1077-1083.
- Bein A, Shin W and others (2018). Microfluidic organ-on-a-chip models of human intestine. *Cellular and Molecular Gastroenterology and Hepatology*, 5(4), pp. 659-668.
- Rajewsky N and others (2020). LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature*, 587(7834), pp. 377-386.

Declarations

Funding

Supported by the Austrian Science Fund (FWF) grant P 35770-B and the German BMBF grant 01ZX2111. Ethics: CETU-EC-2019-016; EIAI-EC-2019-012.

Conflict of Interest

Lukas Garcia and Laura Costa declare no competing interests.

Data Availability Statement

ODRQI scores, scRNA-seq data (GEO GSE238491), and analysis scripts deposited at Zenodo: <https://doi.org/10.5281/zenodo.19542675-data>.

Ethical Approval

Organoid establishment from patient tissue under CETU-EC-2019-016 and EIAI-EC-2019-012, compliant with the Declaration of Helsinki. Written informed consent from all tissue donors.

Appendix A

ODRQI Scoring and AI Pipeline

How we scored each dimension, and how the CNN drug sensitivity pipeline works.

A1. ODRQI Scoring Rules

Structural fidelity (weight 0.284): pathologist histoscore (0-10) times 0.6, plus cell type presence fraction times 0.4.

Disease phenotype (weight 0.264): scored 0-1 against published genotype-phenotype references. CFTR swelling delta-area; tumour WGS concordance; Alzheimer's amyloid IHC.

Drug predictivity (weight 0.184): Spearman correlation of organoid IC50 with clinical response. 0.84 or above = score of 1.0.

Genomic stability (weight 0.148): 1 minus (CNV count at passage 10 divided by 30). Five or fewer CNVs = 0.83+; more than 30 = zero.

Throughput (weight 0.120): wells screened per FTE per week, divided by 384. AI-automated 384-well = 1.0; manual 24-well = 0.063.

A2. Top Ten Programmes

1. CRC-PDTC-AI-14 (colorectal; AI-scored; 384-well): ODRQI 0.924. Phase II AUC 0.884.
2. CF-intestinal-FIS-08 (CFTR swelling assay; EMA-validated): ODRQI 0.904. AUC 0.884.
3. PDAC-PDTC-AI-22 (pancreatic; AI-scored): ODRQI 0.884. Chemo selection concordance 84%.
4. Breast-PDTC-AI-18 (breast cancer; triple-negative subtype): ODRQI 0.864.
5. Hepatic-DILI-28 (drug-induced liver injury screen; 200 compounds): ODRQI 0.844.
6. CF-intestinal-rare-12 (rare CFTR mutation; treatment access): ODRQI 0.824.
7. CRC-PDTC-manual-06 (colorectal; manual scoring; validated): ODRQI 0.804.
8. Lung-COVID-14 (SARS-CoV-2 infection model; antiviral screen): ODRQI 0.784.
9. IBD-intestinal-22 (Crohn's disease; TNF response prediction): ODRQI 0.764.
10. AD-cerebral-18 (Alzheimer's APP mutation; amyloid model): ODRQI 0.744.