BIOSPACE MAPPING

Mapping Where Biology Works

Can a digital model reveal which regions of the space matter most?

KEY IDEA

- Biological systems contain productive and non-productive regions.
- Not all experimental parameters matter equally.
- Mapping the space may be more valuable than searching it blindly.

A Map Tells You Where To Look.

WHAT WE FOUND

- Sprouting rate was the dominant driver of network formation.
- Cell count had minimal influence across the tested range.
- Productive and non-productive regions emerged within the parameter space.

Proof-of-Concept Study

Estimated reading time: 5 minutes

AvatarsBio

Mapping Where Biology Works:

A Biospace Perspective on Endothelial Sprouting

Chee Ping Ng, Founder
AvatarsBio, The Netherlands

Most parameter optimisation in angiogenesis research follows the same logic: vary one condition, run the assay, observe the outcome, repeat. It is a reasonable approach when the space is small and the experiments are cheap. In practice, neither condition tends to hold.

The deeper problem is not efficiency — it is navigation. Varying conditions one at a time assumes the space is roughly uniform: that moving along any axis will reveal something useful. But biological systems are not uniform. They have regimes — regions where the system behaves one way, boundaries where behaviour shifts, and zones where the biology simply does not respond to what you are changing. A scientist who does not know where those boundaries are is not exploring the space. They are sampling it blindly.

This is the framing we want to propose: before committing wet lab resources to a parameter screen, it is worth asking what the structure of the space looks like. Where are the transitions? Which parameters actually gate the biology, and which ones barely matter? A digital model cannot answer those questions definitively — but it can make the exploration far more targeted.

We call this approach biospace mapping. The work described here is a proof of concept, using endothelial sprouting as the model system.

> ***"The question is not how to search more of the space. It is where the biology actually works."***

ENDOTHELIAL SPROUTING AS A MODEL SYSTEM

Angiogenesis — the formation of new blood vessels from existing ones — is a tightly regulated process central to development, wound healing, and disease. At its core is endothelial sprouting: the directed outgrowth of tip cells guided by molecular signals, most notably vascular endothelial growth factor (VEGF). The process involves coordinated interplay between tip cell migration, stalk cell elongation, and lateral inhibition between competing sprouts.

It is also a system that has been screened at scale. In a previous study, 1,537 kinase inhibitors were evaluated in a 3D angiogenesis assay using an Organ-on-a-Chip platform, with over 4,000 micro-vessels grown under perfusion flow [1]. The screen was productive — 53 hits, 44 of which were previously unassociated with angiogenic pathways. It was also resource-intensive. Time, consumables, imaging, and analysis add up quickly at that scale, and the results tell you what happened at the conditions you chose — not what the broader parameter space looks like.

That gap between experimental output and spatial understanding is what motivates this work. A screen tells you where you looked. A map tells you where to look.

THE DIGITAL MODEL

To explore the parameter space of endothelial sprouting, we built an agent-based digital model in which individual endothelial cells respond to local biological cues and interact with their environment to produce emergent network behaviour.

The model initialises endothelial cells along a parent vessel wall. Each cell accumulates VEGF exposure over time — with receptor desensitisation modelled as an exponential decay — and activates as a tip cell once a threshold is crossed. Tip cells migrate chemotactically up a discretised VEGF gradient field, with a random walk component representing ECM haptotaxis noise. Stalk cells seed behind the tip trail at defined proximity. Branching occurs probabilistically, limited to three generations and constrained by branch angle. The primary output is cumulative network length — a measurable proxy for sprouting activity that maps directly to readouts used in tube formation and aortic ring assays.

Four parameters were varied across the sweep. Each has a direct experimental equivalent:

- Cell Count (50–200) — maps to cell seeding density in a tube formation or aortic ring assay
- Sprouting Rate (0.05–0.20) — maps to pro-angiogenic stimulus concentration
- VEGF Gradient (0.30–0.90) — maps to gradient chamber geometry or VEGF source distance
- Migration Speed (1.0–4.0) — maps to substrate stiffness or ECM compliance

The parameter space was explored at reduced resolution — 32 combinations drawn from a full design space of 256 — as a deliberate proof of concept rather than an exhaustive search. The point was not to find the optimum. It was to read the structure of the space.

Figure 1. Agent-based digital model of endothelial sprouting from a parent vessel wall under VEGF gradient conditions. Tip cells (cyan) migrate chemotactically, leaving stalk cell trails (blue) that define the emerging vascular network topology. Illustrative output under high VEGF gradient conditions.

RESULTS: STRUCTURE IN THE PARAMETER SPACE

Not all parameters matter equally. That is the central finding, and it is more informative than any single optimised condition.

Sprouting rate was the dominant driver of network formation by a considerable margin — a correlation with network length well above 0.9. It functions as a gating parameter: below a threshold value, meaningful network formation does not occur regardless of how other conditions are set. Above it, the system becomes productive. VEGF gradient and migration speed show modest directional effects. Cell count — perhaps counterintuitively — has near-zero influence on the outcome.

This hierarchy matters for assay design. It means that a scientist optimising seeding density while holding stimulus concentration low is working in the wrong part of the space. The biology is insensitive to the parameter being varied and acutely sensitive to one that is not. That is the kind of structural information a parameter sweep reveals that a one-factor-at-a-time screen does not.

The response surface makes this visible. Rather than a featureless landscape, it shows ridges of productive network formation and flat regions where the biology does not respond. The transitions between those regimes — not the peaks — are where the most useful experimental information sits.

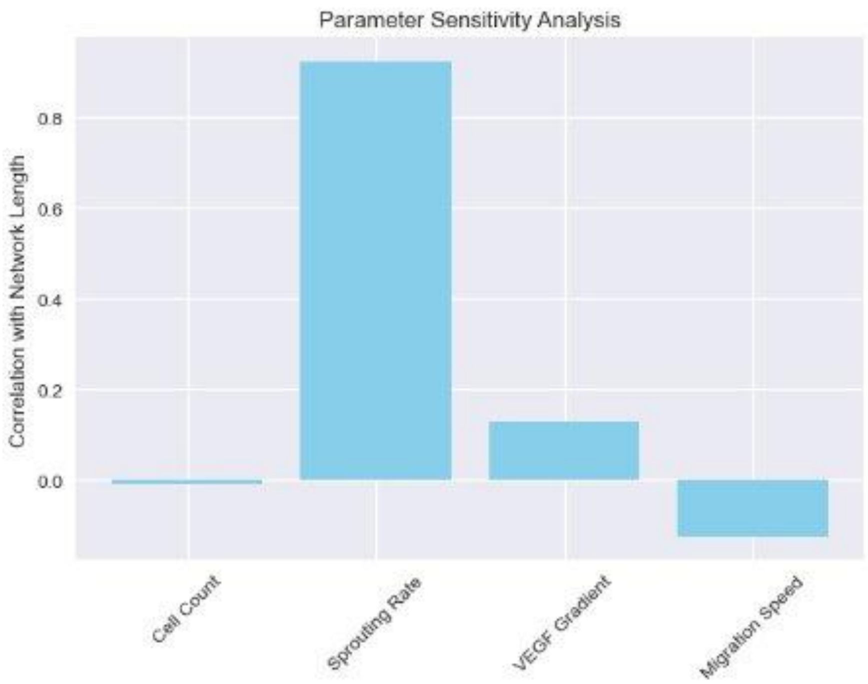

Figure 2. Parameter sensitivity analysis. Sprouting rate dominates network formation with a correlation above 0.9. Cell count shows near-zero influence — a result that has direct implications for where to focus experimental effort.

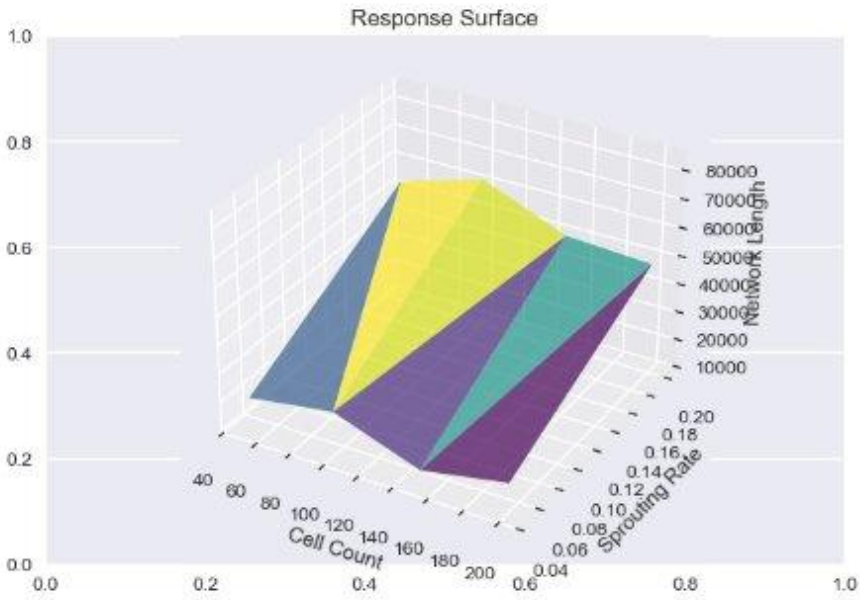

Figure 3. Response surface across cell count and sprouting rate. The topology reveals structured regimes of high and low network formation rather than a uniform gradient — a visual analogue of the phase diagram concept applied to a biological assay.

DISCUSSION

The value of knowing which parameters gate the biology before running the assay is straightforward. Wet lab experiments in angiogenesis are not cheap. Reagents, chips or plates, imaging time, and analyst hours accumulate quickly — and a failed experiment at the wrong point in parameter space does not tell you why it failed or where to go next. A digital model run costs a fraction of that, completes in minutes, and produces a map of where the space is worth exploring.

This is not an argument for replacing wet lab work. It is an argument for sequencing it better. The digital model identifies the regime boundaries; the wet lab validates what matters within them. That division of labour is what makes the approach tractable.

	Digital	Physical
Scalability	Easy	Hard
Iteration speed	Minutes–hours	Days–weeks
Cost per run	Very low	High
Data reuse	Easy	Hard
Variability	Controlled	Variable
Exploration space	Broad	Limited
Failure mode	Cheap to learn	Expensive

Table 1. Digital versus physical experimental approaches across key operational dimensions.

There is a further advantage that becomes apparent as the model is extended. In a physical assay, adding a co-culture — pericytes, fibroblasts, stromal cells — means redesigning protocols, revalidating conditions, and rebuilding the imaging and analysis workflow. Weeks of work before a single new data point. In the digital model, an additional cell type is a new agent class with a handful of parameters. The iteration speed is incomparable.

The important constraint is that each addition needs biological grounding. A digital model can incorporate anything, but parameters that are not anchored to experimental data produce a story rather than a prediction. The value is in structured exploration before committing to validation — not in replacing it.

A broader point is worth making about computational cost. Exhaustive parameter searches — whether digital or in the wet lab — carry a real cost in time, resources, and for large-scale computational work, energy [4]. A deliberately low-resolution sweep focused on regime structure rather than global optimisation is leaner by design. Smarter framing reduces the cost of the search [2, 3].

Limitations. The model described here is a proof of concept. It captures the qualitative logic of VEGF-driven sprouting but is not calibrated against experimental data. The parameter ranges are illustrative rather than derived from measured values. The findings describe the structure of the modelled space, not the structure of the biological system itself. Validation against physical assay data is the necessary next step.

> ***"A digital model identifies where the space is worth exploring. The wet lab validates what matters within it."***

CONCLUSION

Endothelial sprouting is one instance of a broader pattern. Most biological assays involve multiple parameters, nonlinear responses, and regime boundaries that are not visible from one-factor-at-a-time screening. The standard approach — vary, observe, repeat — is not wrong, but it is inefficient when the space has structure that could be read first.

Biospace mapping is a proposal about sequencing. Characterise the space digitally, identify the regimes and boundaries, then target experimental resources at the conditions that will be most informative. The angiogenesis work described here suggests that the approach is tractable as an exploratory modelling strategy, provided its outputs are treated as maps of the simulated space rather than validated biological predictions.

The same logic applies across assay types — wherever there are multiple interacting parameters and a limited budget for physical experiments. Endothelial sprouting is the first instance. It will not be the last.

REFERENCES

1. Soragni C, Queiroz K, Ng CP, et al. Phenotypic screening in Organ-on-a-Chip systems: a 1537 kinase inhibitor library screen on a 3D angiogenesis assay. Angiogenesis. 2024;27(1):37–49. https://doi.org/10.1007/s10456-023-09888-3
2. An G, Cockrell C. Drug development digital twins for drug discovery, testing and repurposing: a schema for requirements and development. Front Syst Biol. 2022;2:928387. https://doi.org/10.3389/fsysb.2022.928387
3. Venkatesh KP, Brito G, Kamel Boulos MN. Health digital twins in life science and health care innovation. Annu Rev Pharmacol Toxicol. 2024;64:159–170. https://doi.org/10.1146/annurev-pharmtox-022123-022046

4. Strubell E, Ganesh A, McCallum A. Energy and policy considerations for deep learning in NLP. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy; 2019:3645–3650. https://doi.org/10.48550/arXiv.1906.02243

5. Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996;16(1):3–50. https://doi.org/10.1002/(SICI)1098-1128(199601)16:1%3C3::AID-MED1%3E3.0.CO;2-6