0.1 Background

Tumour spheroids (and organoids) are 3D culture models that bridge the gap between 2D cell lines and animal studies by providing genetically and histologically representative tumour architecture while preserving cellular heterogeneity [?]. They reproduce proliferative, quiescent and necrotic zones, making them strong candidates for in-vitro drug screening. Although early spheroid methods were labour-intensive, several automated platforms now enable high-throughput generation of uniform spheroids [?] [?] [?] [?].

Bayesian optimization (BO) is a concrete algorithmic application of the Bayesian Decision Theory (BDT), which provide a formal framework for optimal decision making under uncertainty and with prior knowledge (a prior). BDT has been repeatedly proposed to be medical dose finding trials [?], [?], [?], [?]; but has been limited to Bayesian dosing dashboards for inflammatory bowel disease [?], [?], [?].

Single-dimension CGGs are widely used in drug screening [? ? ?]. In contrast, only three distinct designs capable of generating surfaces of concentration gradients have been reported. Two are impractical due to excessive spatial and fabrication complexity, while the other is directly relevant and is incorporated into the proposed research plan[?] [?] [?].

BO, originally popularised for ML hyperparameter tuning, has now been adopted in empirical sciences, mainly in context chemistry and materials science [? ? ? ?]. Biomedical applications include in-vitro drug-lead screening [? ?], protein-formulation optimisation [?], and fitting complex biological models [? ?]. Notably, Yakavets et al. (2025) applied BO to a three-drug sequential regime in breast-cancer spheroids but failed due to poorly encoded variables and an over-dimensional (7D) design relative to sample size (n = 40) [?].

0.2 Research aims and objectives

Hypothesis. A 3-inlet concentration gradient generator microfluidic chip can be paired with physiological-flow replicating tumours spheroid arrays to enable high throughput screening of combinatorial therapies.

Deliverables Objective 1 (O1). Design, simulate and fabricate a microfluidic chip that exposes hydrogel-embedded tumour spheroid arrays to a reproducible 3-solution concentration gradient.

Objective 2 (O2). Demonstrate reliable NSCLC (A549) spheroid formation, viability and drug-response readouts on-chip, and show that known synergistic combinations can be replicated.

Objective 3 (O3). Develop and benchmark a structured-batch Bayesian optimisation (SBBO) workflow tailored to gradient-generated batches, first in-silico and then on the microfluidic platform.

Objective 4 (O4, reach objective). Apply SBBO to explore and identify previously untested candidate synergistic drug sequences for NSCLC, providing ranked combinations for future validation.
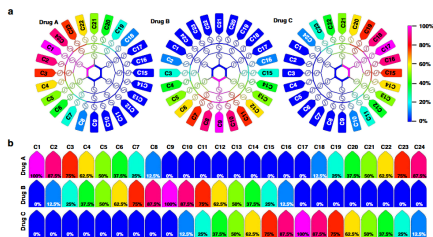
0.3 Research plan

Figure 1: 3-inlet spiral mixer generates 24 distinct linear concentrations. Adapted from Shen et al. [?]
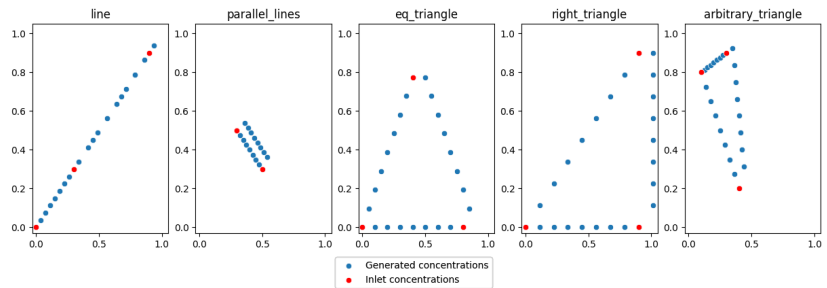
Figure 2: Putative structured batch shapes within-two drug dose-space. Each batch is defined by three inlet concentrations and corresponds with a projection from a higher-dimensional drug space.

1

1. **Design and Build a 2D MGG chip platform for Tumour Spheroid screening (O1)**: The chip will be developed using SolidWorks and COMSOL CFD, with fabrication via standard PDMS photo-lithography workflows available in the CRUK nanofabrication facility [?]. These established tools allow rapid design–test iterations, ensuring feasibility within the project timescale.

   a) **Design a modified 3-solution mixer** Although PDMS chips can be fabricated reliably via photolithography and soft lithography, the primary throughput bottleneck remains the limited number of syringe pumps available for delivering distinct conditions

      i. **Expand and test the 3 solution mixer** Thus, we adapt the concentric ring Tai Chi spiral 3-inlet mixer from Shen et al, which generates 24 linear combinations from 3 injected solutions by [?]. [?]. Channel width and geometry will be adjusted to support flow of $96\,\mu\mathrm{m\,s^{-1}}$ needed for the larger spheroid arrays rather than the original single-cell screening array design. The mixer is designed following recursive pattern - generating $3 * 2^n$ unique concentrations for n. Thus, if CFD predicts stable flow, scaling beyond 24 outputs (e.g., to 48–96) will be attempted; otherwise the validated ($R^2 = 0.9946$) 24-output design from Shen et al. will be used [?] Figure 1.

      ii. **Test mixing robustness**. Mixing performance and noise will be quantified by perfusing fluorescein through one inlet and water through the other two, followed by fluorescent imaging. If acceptable concentration noise levels are not achieved (concentration Signal to Noise (StN) ratio > 5%), the number of mixer rings will be decreased. Since all inlets are supposed to inject fluid at identical rate, a design where single motor drives 3 syringes can be implemented.

   b) **Integrate a physiological flow Tumour spheroid array** Prince et al (2022) describes a well-scaling technique to generate Tumour Spheroid arrays under near-physiological flow conditions [?].

      i. **Confirm Tai-Chi spiral comparability with hydrogel injection**. This hydrogel-loading method already works in "Christmas tree" CGGs, so similar behaviour is expected in Tai-Chi mixers. We will test a configuration of 8 spiral stages feeding five 20-cavity arrays to confirm that EKGel precursor consistently fills the cavities, first in CFD and then experimentally. EKGel (1% a-CNCs and 2% gelatin) is employed due to its long gelation time of 50 minutes, which gives sufficient time to invade the cavities and can be purged from channels using fluorinated oil [?]. Afterwards, the whole setup will be investigated by brightfield microscopy. Lack of hydrogel within tumour-designated cavities and presence within mixing system will be considered a failure. If cavity filling is inconsistent or hydrogel cannot be purged from the mixing system, a dedicated phase-guide or micro-post separated channels will be introduced. These can later be sealed with adhesive tape or epoxy glue [?].

   c) **Integrate 3-solution mixer with cavity arrays**. The cavity arrays (4 lines, 20 each) will then be integrated at the end of of the 3-solution mixer. Upon successful hydrogel-cavity filling, the CGG capabilities (fluorescence inside cavities) will be reevaluated as in step [step Test Mixing Robustness]. If successful, hydrogel, fluorinated oil, and media injection through single inlet (remaining inlets plugged) will be investigated.

2. **Benchmark drug-dose response generation (O2)** We next benchmark the device's ability to produce reproducible drug–response data. Quantifying variance in delivered concentrations and spheroid viability is essential for subsequent BO modelling.

   a) **Spheroid Generation**. The seeding approach from Prince et al is adapted (cancer cells premixed with EKG precursor [?]) using the A549 cell line derived from a human lung adenocarcinoma, a non-small cell lung cancer (NSCLC). The A549 can be acquired from the ECACC COSMIC cell lines, available through the UK Health and Security Agency [?]. A549 spheroids reliably form in matrices similar to EkGEL (e.g. Matrigel or polyPEGDA–gelatin) [?]. EKGel has similar properties and is used due to slow-gelling time required for cavity filling procedure [?], [?]. Prior to hydrogel mixing, the cells will be incubated with 20 µM of Cell Explorer Live Cell Tracking dye (or alternative) and resuspended in media, to allow Two-photon Confocal Laser Microscopy Scanning (TP-CLSM) [?]. Spheroids will mature over 5 days under continuous perfusion, following established protocols [?].

   b) **Linear-concentration benchmarking**. Gradient performance will be validated by 5 day culture with cisplatin, which produces consistent spheroid toxicity profiles [?]. Three inlet concentrations (0, 5, 10 µM) spanning the active range [?] yield 24 discrete conditions. This provides a feasible, low-cost cell-viability benchmark for noise levels and later multi-drug tests.

   c) **Integration of automated imaging analysis** Jeon et al 2025 TP-CLSM analysis pipeline will be

adapted to spheroid arrays, which have lower spatial complexity than inverted colloidal crystals, minimising development time.[?]

d) **3 Drug benchmarking**. Three-drug benchmarking will use the same protocol as the cisplatin gradient, but with fixed inlet concentrations for two additional agents: necitumumab and gemcitabine. This restricts experimental scope to synergistically confirmed 3-drug combination therapy [?] while demonstrating compatibility with multi-drug designs. Moreover, the cell viability measurements from a single chip are sufficient for estimation of effect-based drug synergy metrics (e.g. bliss-independence) [?]. Moreover, the chips suitability to generate linear concentration gradients, readily allows to adapt it to more robust dose-effect models [?], [?].

e) **Publication** Regardless of BO integration, a validated novel spheroid–gradient platform is a publishable outcome (e.g., *Nano Letters*, *Advanced Healthcare Materials*, *Microsystems & Nanoengineering*, *Lab on a Chip*). This ensures intermediate deliverables even if later milestones prove challenging.

3. **Design and in-silico test a experimental design tailored Bayesian Optimisation (O3)**. Standard batch BO assumes free target selection, whereas CGGs yield structured concentration sets. Since physical and labour microfluidic screening constraints are harder to solve, an active-learning policy must thus be adapted to work under these constrains.

a) **Implement Structured-Batch Bayesian Optimisation (SBBO) algorithm**. As shown in Figure 2, the three inlet concentrations allow exploring the parameter space along discrete points on edges of arbitrary triangles.

    i. **In-silico testing setup** Note that *This work is purely computational and will run during experimental downtime*. BO will be implemented in Python using BoTorch/PyTorch libraries [?], [?]. All algorithms will be benchmarked using Ackley, Rosenbrock, and Hartmann functions within relevant dimensionality space [?]. Evaluation metrics will be: 90% convergence speed, simple and cumulative regret [?].

    ii. **Implement a linear-space acquisition function**. An acquisition function defines the salience of any data point being investigated. We will adapt the acquisition function to score lines of n discrete points between two inlets by cumulative (all points on the line) utility, and validate convergence to an optimum on 3D test functions.

    iii. **Implement a structured batch acquisition function** We then extend the paradigm to structured triangular batches defined by three inlet concentrations Figure 2. If the triangle search space becomes too large, we will use a secondary optimisation routine to propose candidate triangles.

b) **Benchmark and computationally optimize SBBO**

    i. **Low-dimension benchmarking** SBBO will be benchmarked against batch-BO, evolutionary methods, binary search, random, and grid search in 4D. Each algorithm evaluates up to 216 points across 100 seeds. SBBO should underperform batch-BO but exceed baseline methods. The simulations will be ran in parallel on Imperial's HPC systems.

    ii. **Algorithm optimization (reach objective)** The unoptimized algorithm will likely be computationally expensive in high dimensions, thus making less myopic approaches unfeasible. Standard BO optimisation tricks will be implemented and any remaining downtime spent on solution search [?]. Given the days-long duration of empirical microfluidic batch, hours-long runtime is acceptable.

    iii. **High-Dimensional Benchmarking (reach objective)** SBBO will be tested in 6D using 7D benchmarks, mimicking six-drug synergy searches. It should outperform all baselines except BO variants. If BO is unexpectedly outperformed by an alternative algorithm capable of structured batch generation, it will be abandoned in favour of the algorithm.

c) **Adapt SBBO for biological noise** Once empirical platforms noise estimates are available, SBBO will be adapted for heteroscedastic settings.

    i. **Adapt objective function to noise**. Objective functions model the underlying distribution constructed from investigated points and are used when generating new targets [?]. To avoid noise-driven overfitting, empirical noise levels will be incorporated into the objective by use of Matern kernel, homoscedastic or heteroscedastic gaussian process, or more suitable alternative.

    ii. **Benchmark SBBO in noisy environment** The algorithm capability to discover optimums will be reevaluated in noisy settings. The exploration efficiency is likely to drop; inability to rediscover an optimum or optimum-proximate region will be considered an objective failure of the algorithm. However, this needs empirical estimation of platforms noise levels.

4. **Run empirical SSBO trial (O4)** A proof-of-concept SSBO-guided campaign will be run in a four-drug space. Exploration of five to six drugs will be treated as a reach objective, pursued only if earlier milestones are achieved ahead of schedule. At the appropriate stage, the system will be used to identify promising drug candidates in accordance with future research trends.

    a) **Establish single-drug dose–response curves.** Single-drug gradients (one chip per drug, run in parallel) will be used to fit simple dose–response models. These will define the initial prior and objective for SSBO, reusing the linear-gradient protocol from O2 to keep the experimental workload manageable.

    b) **Lab-in-a-loop workflow.** SSBO will be run for **up to 10 batches** (two chips per batch), which is sufficient to demonstrate closed-loop operation without exceeding realistic chip and culture throughput. For each batch, SSBO proposes three inlet drug cocktails per chip; after imaging, spheroid viability metrics are extracted and fed back into SSBO to generate the next batch.

    c) **Results analysis.** We will evaluate SSBO identified optimums against control and known combinatorial studies. Evidence that SSBO achieves better viability outcomes or finds promising regions faster will demonstrate practical value, even without fully mapping the search space, and would constitute a publishable case study.

0.4 Pathway to Impact

Multi-drug treatment dose–response relationships are currently underexplored compared to their opportunity to improve treatment outcomes [?] [?]. Experimentation is costly in physiologically relevant models and limited by both hardware and experimental-design constraints. Microfluidic chips used in screening investigate single concentrations only or generate, at most, a few linear concentration gradients, which inflates experimental footprint [?] [?][?] [?]. Typical workflows are both equipment and labour intensive, requiring multiple experiments to investigate dose-response synergies of just two drugs.

By explicitly coupling a three-inlet gradient generator with tumour spheroid arrays and a structured-batch Bayesian optimisation algorithm Figure 2, this project will deliver a proof-of-concept platform for data-driven exploration of multi-drug therapies. In the short term, the platform reduces the experimental burden associated with identifying promising sequential drug treatments for NSCLC, while providing a reusable framework applicable to other cancer types and drug classes. The system will be built using widely adopted materials (PDMS, gelatin-based hydrogels) and open-source analysis pipelines, supporting reproducibility and transfer learning.

In the longer term, the primary impact lies in precision oncology. Once a baseline SBBO model has been trained on generic cell lines, the same framework can be adapted to patient-derived cells, using transfer learning to rapidly refine treatment sequences under constraints of limited sample material and time. Incorporating prior knowledge into Bayesian optimisation mitigates the costly cold-start phase [?], making this approach particularly suitable for clinical settings. This aligns with current interest in functional precision medicine and provides a principled experimental-design layer on top of existing ex-vivo drug-screening assays. Beyond cancer, the structured-batch optimisation concept is broadly applicable to microfluidic gradient platforms in drug delivery, toxicology, and tissue engineering.