

WP2-1

Breast Cancer / CRC Virtual Cell Model Training, PoC and Technology Transfer Outsourcing Specification

Using the WP1 Model-ready Databank to establish Drug Response, Genetic Perturbation, and Cross-context Drug Perturbation capabilities

- Client: ITRI
- Contractor: Amrita Chattopadhyay, Assistant Professor, Global Health Program (GHP) / Institute of Food Safety and Health (IFSH), College of Public Health, National Taiwan University (NTU)
- Contract Amount: NTD 240,000 total
- Document Date: 2026-09-22
- Final Delivery Date: 2026-11-13
- Final Acceptance Date: 2026-11-20

1. Background and Scope

WP1 has completed the data collection and Model-ready Databank. WP2-1 will use the available WP1 data as the primary input to implement, evaluate, demonstrate, package, and transfer three Virtual Cell model capabilities.

Completion will be based mainly on executability, reproducibility, and handover readiness. Breast cancer is the primary application domain. CRC is the secondary evaluation domain and will be included where WP1 provides compatible data; CRC-specific perturbation PoC is not required for every model.

- WP2-1 focuses on reproducing, adapting, and applying existing public model architectures. Development of a new large-scale model from scratch is not required.
- All PoCs are for research use only and are not patient-specific clinical treatment recommendations.

2. Objectives

1. M1 - Drug Response / IC50: Use baseline expression / Cell State and drug information to predict a D1-agreed quantitative drug-response endpoint and rank drugs within the same cell context. IC50 shall be reported where it can be reliably estimated from dose-response data, and LC50 shall be reported only where the assay measures lethality and supports a reliable estimate.
2. Disease PoC: Complete demonstrable Breast cancer drug-response and perturbation examples, plus at least one CRC evaluation where compatible data are available.
3. Packaging & Transfer: Deliver code, model weights, environment definitions, examples, documentation, and handover materials so that ITRI can run the main inference workflows independently.

3. Models and Minimum Work Scope

Model	Reference approach	Minimum input	Minimum output	Minimum requirement
M1 Drug Response / IC50	PERCEPTION-style or equivalent agreed by both parties	Expression / Cell State + Drug	D1-agreed quantitative drug response endpoint and drug ranking; IC50 / LC50 where scientifically supported (Section 5.1)	At least one simple baseline and one commonly used ML baseline

Only code, model weights, and data with terms permitting the agreed commercial use and the required handover to ITRI may be included in the deliverable workflow. Non-commercial-only materials are excluded. An alternative to a designated model or dataset requires written approval by ITRI and a corresponding update to D1.

4. Common Execution Principles

- Data source: Use compatible WP1 data with documented rights for commercial use and redistribution to ITRI. Record the original source, exact release, license, required notices, and intended use for

every dataset. Appendix A lists eligible releases and restrictions. Public availability or a license on a repackaged copy alone does not establish the original redistribution rights.

- **Multi-dataset integration:** Where at least two independent, compatible datasets are available and confirmed in D1, each applicable experiment series shall compare separate-dataset training, simple pooling, and at least one method that accounts for source differences using a jointly trained model. Use common split manifests and the same held-out evaluation samples and metrics; each separate-dataset model uses only its own training-source subset. D1 shall specify the shared features/endpoints and the biological differences to preserve.
- **Integration design:** Align common genes, drugs, and task definitions while retaining dataset, specimen, donor/cell-line, batch, and assay identifiers. Different donors remain separate observations; matching local patient IDs across studies is not evidence of the same person. Joining complementary tables or splitting one dataset into files does not by itself count as multi-dataset fusion.
- **Source sufficiency:** D1 shall resolve whether each task has enough compatible, commercially transferable datasets. If not, the parties must approve an alternative source or an explicit scope amendment in writing. A single-dataset result cannot be labeled as completion of the multi-dataset integration requirement.
- **Data split:** D1 shall define the held-out unit for each task, such as patient, cell line, gene perturbation, drug, or dataset. Records belonging to that held-out entity, including duplicates across datasets, shall remain in the same split, except for target-context controls explicitly permitted by the D1 adaptation protocol. Fit feature selection, normalization, batch correction, and integration on training data only; use validation data for model selection. Test responses shall not be used for fitting, tuning, or choosing the integration method. Any target-context control adaptation must be specified in D1 and labeled as described in Section 5.3.
- **Evaluation:** Each model will use primary metrics, baselines, and repeated-run settings agreed in D1. Report individual runs, the aggregate result, and variability under the same splits and evaluation samples. Include a simple statistical baseline, which may be competitive with complex models. A single numerical score will not be the sole basis for acceptance.
- **Reproducibility:** Keep the main configs, checkpoints, data versions, split manifests, preprocessing and integration parameters, and execution instructions so that ITRI can rerun the main results. Deliver the integrated model-ready table where redistribution is permitted; otherwise, provide a reproducible build procedure from the approved data, with source IDs and mappings.
- **Result handling:** If model performance is below the initial expectation, the actual result, a brief analysis, and applicable limitations shall be delivered. No fixed numerical performance threshold is required unless separately agreed in writing.
- **Transfer licensing:** The commercial-use requirement applies to data used to train and evaluate the deliverable models as well as the files handed over. Check code and model-weight terms separately. Delivering weights alone does not resolve unclear rights in their training sources. Preserve required attribution, license notices, modification records, and any applicable share-alike obligations.

5. Model Specifications

5.1 M1 Drug Response / IC50

- **Data:** Use DepMap Public 24Q4 expression data and PRISM Repurposing 19Q4 drug-response data under CC BY 4.0. Where an additional compatible, commercially transferable response dataset is available and confirmed in D1, it shall be included in the multi-dataset comparison. Joining DepMap expression to PRISM labels alone does not satisfy that comparison. GDSC and CTRPv2 require documented permission or rights clarification before inclusion in training, evaluation, or transfer.

- **Evaluation:** Include at least one held-out evaluation and show predicted ranking of multiple drugs within the same cell context. Where at least two compatible datasets are confirmed in D1, compare separate-dataset, pooled, and source-adjusted joint training on the same evaluation cases. D1 shall define compatible response endpoints, units, and dose windows; incompatible assay endpoints shall not be pooled as interchangeable labels.
- **Metrics:** Report at least one correlation or error metric and one ranking metric; the exact metrics will be confirmed in D1.
- **Outputs:** predicted response, D1-agreed quantitative response endpoint and unit, rank, drug/cell ID, model version, and observed response when available for evaluation. AUC shall be reported where supported by source assay. Provide IC50 only where it can be reliably estimated from the dose-response curve, and LC50 only where the assay measures lethality. Flag estimates outside the tested concentration range and values that cannot be reliably estimated; do not relabel IC50 or viability as LC50.

6. Disease PoC

- **Breast PoC-1:** At least one Drug Response / ranking example showing input, model output, observed response where available, and a short interpretation.
- **Breast PoC-2:** At least one gene or drug perturbation example showing Δ RNA, Top DEG, and pathway/state shift.
- **CRC Evaluation:** At least one CRC evaluation case using the task best supported by compatible WP1 data; M1 is acceptable.
- If direct Breast/CRC perturbation ground truth is unavailable, the result must be labeled as model inference or transfer evaluation and not as validated patient treatment response.

7. Deliverables and Simplified Acceptance Criteria

ID	Deliverable	Minimum content	Acceptance criteria
WP2-D1	Task Definition & Evaluation Plan	Tasks and endpoints; approved data releases and licenses; model/reproduction scope; multi-dataset integration comparisons where at least two compatible datasets are available and confirmed in D1; split/control-adaptation rules; metrics, baselines, repeated-run settings, and environment for all three models	ITRI confirms the scope and evaluation plan in writing, including identification of at least two compatible datasets for each applicable joint-training comparison. Where two compatible datasets are not available, D1 shall document the reason and define an alternative cross-dataset evaluation or scope amendment, subject to prior written approval by ITRI.

ID	Deliverable	Minimum content	Acceptance criteria
WP2-D2	M1 Model Package	M1 code, weights/checkpoint, example, main results, integration comparison results, where applicable under the D1-agreed multi-dataset scope, and brief notes	ITRI can run the designated example from README and obtain drug-response/ranking output.
WP2-D3	M2 Model Package	D1-agreed GEARS reproduction, training/inference code, configs, weights, example, integration comparison results, where applicable under the D1-agreed multi-dataset scope, and DEG/pathway outputs	ITRI can run inference for a designated perturbation and obtain post-RNA/ Δ RNA and DEG/pathway results.
WP2-D4	M3 Model Package	D1-agreed CRISP-style/equivalent code, weights, held-out-context example, control-adaptation setting, and integration comparison results, where applicable under the D1-agreed multi-dataset scope	ITRI can run inference for the designated strict held-out context without target-context model adaptation. Control-assisted transfer, if evaluated, shall be reported separately.; direct comparison is provided when ground truth is available, with the integration comparison results supplied where applicable under D1.

8. Overall Acceptance Principles

- The main acceptance basis is successful execution, complete outputs, clear documentation, and reproducibility of the main workflow, including the D1-agreed multi-dataset training and comparison results, where applicable under the D1-confirmed compatible-dataset scope. A joint table without joint training, or an inference demonstration without available ground-truth evaluation, shall not be described as demonstrated integration benefit.
- Acceptance does not require the model to outperform every comparison method. Actual evaluation results shall be reported as obtained.
- Small numerical differences caused by random seed, hardware, or package versions are acceptable when they do not change the main conclusion and are documented.
- CRC work depends on compatible WP1 data. The absence of compatible CRC data for a specific model is not, by itself, a rejection condition.

For any mandatory item that cannot be completed because of a documented technical constraint, any alternative deliverable or scope amendment requires prior written approval by ITRI. A limitation statement alone does not satisfy acceptance of a mandatory deliverable.9. Delivery, Review, and Acceptance Dates

Delivery date	Deliverables	Acceptance completed by	Acceptance focus
2026-10-09	D1: Task Definition & Evaluation Plan	2026-10-14	Confirm tasks, data, splits, metrics, and model choices
2026-11-06	D2-D4 Beta: runnable versions of all three models	2026-11-13	Confirm all three models run and produce the main examples and initial results

Schedule rule: The dates above are the formal WP2-1 delivery and acceptance checkpoints. Minor corrections, if needed, should be completed before the corresponding acceptance date. Final acceptance shall not be later than 2026-11-13.

10. Explicit Exclusions

- No requirement to build a new large general-purpose Virtual Cell model from scratch.
- No clinical-grade patient-specific drug recommendation, medical-device validation, or clinical trial work.
- No requirement to use non-public or internal sensitive data unless separately agreed by both parties.
- No guarantee that a model must outperform every comparison method. The contracted scope is reproducible model implementation, evaluation, PoC, packaging, and technology transfer.
- PoC, packaging, and technology transfer are included in WP2 and are not separate work packages.

11. Contract Amount

Item	Amount	Main scope
M1 Drug Response / IC50	NTD 240,000	Model implementation, evaluation, ranking, packaging, and transfer

Total WP2-1 contract amount: NTD 240,000. Disease PoC, Unified Packaging, Technical Documentation, and Technology Transfer are included in the scope above.

Approval / Confirmation

ITRI Project Lead: _____ ITRI Technical Acceptance: _____

Contractor Project Lead: _____

Appendix A. Dataset releases for D1 selection

- This list identifies candidate releases with published licenses information that appears to permit the agreed commercial use and redistribution, subject to their applicable conditions and D1 verification. It is a selection list, not a requirement to use every source or a claim that the sources are scientifically interchangeable. D1 shall record the exact files, version/hash, intended task, and applicable license evidence. Final eligibility for use and transfer shall be confirmed in D1 based on the exact release, original-source terms, and intended use.
- DepMap Public 24Q4: Figshare item 27993248, version 1; CC BY 4.0. Expression features for M1. This does not cover every CCLE or DepMap release.
- PRISM Repurposing 19Q4: Figshare item 9393293, version 4; CC BY 4.0. M1 drug-response data, including secondary dose-response measurements.
- Replogle 2022: Figshare item 20029387, version 1; CC BY 4.0. K562 genome-wide, K562 essential, and RPE1 essential processed Perturb-seq panels for the D1-selected M2 benchmark scope.
- VCC2025: Arc Virtual Cell Challenge 2025 released H1 hESC training, validation, and test processed datasets; CC0 1.0. Genetic-perturbation benchmark candidate; not Breast or CRC data.
- Jiang dataset: Zenodo record 14518762, version 2.1; CC BY 4.0. Stimulus-context perturbation data; use for M2 only after confirming the perturbation type and target mapping.
- MIX-seq: Figshare item 10298696, version 3; CC BY 4.0. M3 candidate with 13 drugs, a single tested dose per drug, and predominantly 24-hour observations. D1 shall limit claims to the selected files and observed coverage.
- Additional integration candidates: KOLF2.1J CRISPRi atlas, Figshare item 27261219, version 1 (human induced pluripotent stem cells), and DepMap Perturb-seq pilot, item 33273600, version 1; both CC BY 4.0. Their use depends on task compatibility; KOLF2.1J is not CRC data.
- METABRIC, cBioPortal brca_metabric distribution: ODbL 1.0. Commercial use is permitted with attribution and the applicable share-alike/open-database obligations. D1 must define whether an integrated database is a derivative and how its distribution meets those obligations; unrestricted proprietary redistribution shall not be assumed.
- CC BY 4.0 sources require attribution, a license notice/link, and identification of modifications. Non-commercial-only sources are excluded. GDSC, CTRPv2, sci-Plex3, and Norman/Adamson/Dixit copies packaged for GEARS are outside the committed list unless original-source rights or separate permissions covering the required commercial use and redistribution are documented. An open license on a third-party repackaging alone is insufficient.