Peer-Review

Sun, Cheng M. 2026. "*In Silico* Optimization of Time-Dependent Motility and Proliferation Control for Wound Closure: A Hybrid Fisher–KPP and Force-Based Agent Model." *Journal of High School Science* 10 (2): 140–58. https://doi.org/10.64336/001c.161295.

Your key results are modeled by construction- not discovery. It is a mathematical inevitibility of KP dynamics and force-based migration. Second, you are not modeling acoustic healing (in fact you could be modeling any form of healing) because it is only represented as a parameter change; not explicitly modeled as a physical phenomenon. The change in parameter may as well be due to growth factor therapy, immune modulation, laser therapy, electric field…….. You do not represent or model wave propagation (frequency, amplitude, wavelength), mechanical stress/strain fields, mechanotransduction pathways (e.g., integrins, ion channels) and fluid–structure interaction. This is purely phenomenological parameter scaling, not acoustic modeling. Your model is non-identifiable non-falsifiable with respect to acoustics. You cannot test if frequency matters, if amplitude matters and if your effects are consistent with known mechanotransduction.

There is no fitting to experimental wound-healing data, no validation against acoustic stimulation studies, no comparison to baseline Fisher–KPP fits.

SUGGESTIONS: Leave acoustic healing entirely out of the equation. Specific concerns appear below.

1.If you leave acoustic healing entirely out of your paper and reframe your title and content around something like "A hybrid PDE–ABM framework for studying how time-dependent modulation of cell motility and proliferation affects wound closure dynamics", then your paper actually becomes much more defensible and contributes to the literature.

2.Explicitly state that the stimulus is an abstract, dimensionless control that modulates effective diffusivity (D), proliferation rate (ρ), motility force (F_ext). This should be framed as a phenomenological input not tied to any specific therapy. Purge all acoustic related text in the manuscript.

3.Add generality checks:

A. Parameter robustness: Does the threshold persist across:different adhesion strengths? different baseline D, ρ?

B. Mechanism separation: Vary D only, ρ only, and force only. Show which mechanism drives the effect

C. Scaling analysis:Non-dimensionalize:show dependence on key ratios (e.g., motility vs adhesion)

4.Add validation: Fit baseline model to published data (e.g., Fisher–KPP fits).

5.Add explicit limitations: Stimulus is not mechanistically grounded, No modality-specific predictions, No biochemical signaling included. No experimental validation.

6.Quantify stochastic variability: For every reported metric, such as wound area vs time, closure time, % closure at 48 h, run $N \geq 20–50$ independent simulations with different seeds. Report Mean trajectory, Variance (or SD / SEM), Confidence intervals (preferably bootstrap). Without uncertainty quantification, your phase diagram is not trustworthy. The 'threshold' claim is especially sensitive.

7.Report convergence checks: Does increasing number of cells / grid resolution change results? Report Sensitivity to randomness: variance vs parameter regime: does noise increase near threshold?

8.Code release: Provide 1. Full repository at GitHub of simulation code, parameter files, plotting scripts

Reproducibility package:script to reproduce:Fig. 1 (time series), Fig. 2 (velocity field), Fig. 3 (heatmap)

Seed control: explicit random seed parameter, documented RNG usage.

9.These revisions will actually make the manuscript designable for a very high level impact. You can take the baseline model, and fit it to (actual data of) poor-wound healers (diabetics, hypertensives, immune deficient) from the literature. Then make minimum justifiable changes to

motility or adhesion (one at a time) and fit this to good healers (healthy population). For this, you will need literature real data. Fit to time series, not just endpoints. Fit to one dataset and predict another within the same (good or poor healer regime). If you find that Poor healers → below threshold and Good healers → above threshold, then your threshold becomes biologically meaningful, not just an abstract model number.
You could then state that "Interventions may act by shifting tissue dynamics across a critical motility–duration threshold." This gives researchers a real tangible research-trajectory to work on and makes a significant clinical impact.

_____

Response to the Reviewers

We sincerely thank the reviewers for their careful reading of our manuscript and for their constructive and insightful comments. We have revised the manuscript extensively in response. In particular, we have reframed the study to remove modality-specific acoustic claims, clarified that the stimulus is a phenomenological control input, added stochastic uncertainty quantification, introduced mechanism-separation and robustness analyses, added baseline Fisher–KPP calibration against literature-consistent scratch-assay data, and expanded the limitations and acknowledgments sections. We believe these revisions have substantially improved the rigor, clarity, and interpretability of the manuscript.
For ease of review, all major revisions have been highlighted in the revised manuscript.

Reviewer 1

Comment 1.

Your key results are modeled by construction—not discovery. It is a mathematical inevitability of KP dynamics and force-based migration. Second, you are not modeling acoustic healing … This is purely phenomenological parameter scaling, not acoustic modeling. Your model is non-identifiable/non-falsifiable with respect to acoustics …

Response:

We thank the reviewer for this important conceptual observation. We agree that the original manuscript overstated modality specificity. In response, we have **fully removed the acoustics-specific framing** from the title, abstract, introduction, discussion, and conclusions, and we now explicitly describe the stimulus as a **generic, abstract, dimensionless control input** that modulates effective motility, proliferation, and protrusive forcing. The revised title is now:
In Silico Optimization of Time-Dependent Motility and Proliferation Control for Wound Closure: A Hybrid Fisher–KPP and Force-Based Agent Model.
We also revised the abstract and introduction to state clearly that the model is **not intended to represent acoustic wave propagation or any specific therapy modality**, but rather to study general dynamical principles governing wound closure under time-dependent control. We further added an explicit **Limitations** paragraph in the Discussion stating that the model does not include wave propagation, stress/strain fields, fluid–structure interaction, or mechanotransduction pathways, and therefore does not generate modality-specific predictions.

Comment 2.

There is no fitting to experimental wound-healing data, no validation against acoustic stimulation studies, no comparison to baseline Fisher–KPP fits.

Response:

We appreciate this point and have addressed it by adding a **baseline Fisher–KPP calibration section** to the Results. Specifically, we now include a new subsection, **"3.5 Baseline Fisher–KPP calibration,"** together with a new figure showing calibration of the Fisher–KPP baseline against literature scratch-assay control data. This addition provides a validation-oriented reference point for the hybrid simulations and grounds the baseline dynamics in published wound-healing data.
Because the manuscript is no longer framed as an acoustics paper, we did not pursue modality-specific validation against acoustic stimulation studies; instead, we focused validation on the

baseline wound-healing model itself, which we believe is the more defensible approach in light of the reviewer's conceptual concern.

Comment 3.
SUGGESTIONS: Leave acoustic healing entirely out … reframe around a hybrid PDE–ABM framework for studying how time-dependent modulation of cell motility and proliferation affects wound closure dynamics.
Response:
We fully adopted this recommendation. The manuscript has been reframed accordingly. Acoustic-specific language, examples, and implications were removed throughout, and the central contribution is now presented as a **hybrid PDE–ABM framework for studying wound closure under time-dependent modulation of effective motility, proliferation, and protrusive forcing**.

Comment 4.
Explicitly state that the stimulus is an abstract, dimensionless control that modulates effective diffusivity (D), proliferation rate (ρ), motility force (F_ext). This should be framed as a phenomenological input not tied to any specific therapy. Purge all acoustic related text in the manuscript.
Response:
Thank you. We implemented this suggestion directly. The abstract now states that the control signal is an **abstract, dimensionless input** that modulates effective motility, proliferation rate, and/or protrusive forcing. The introduction and methods sections were revised consistently, and acoustic-specific text has been removed throughout the manuscript.

Comment 5A.
Add generality checks: Parameter robustness: Does the threshold persist across different adhesion strengths? different baseline D, ρ?
Response:
We thank the reviewer for this suggestion. We added a new Results subsection, **"3.4 Threshold robustness,"** together with a new figure showing that the amplitude–duration threshold shifts systematically across different adhesion/drag conditions. This directly addresses the robustness of the threshold claim and supports the interpretation that stronger adhesive resistance requires stronger or longer control to achieve reliable closure.
We note that the present revision explicitly includes robustness across adhesion strength. Baseline parameter sensitivity is discussed in the Methods and Discussion as part of the broader interpretation of the threshold structure.

Comment 5B.
Mechanism separation: Vary D only, ρ only, and force only. Show which mechanism drives the effect.
Response:
We agree that this is essential. We therefore added a new subsection, **"3.3 Mechanism separation,"** together with a new figure showing separate simulations in which only diffusivity, only proliferation, or only protrusive forcing is modulated. These results show that motility-related control pathways dominate closure acceleration on the 48 h timescale, whereas proliferation alone has only a modest effect. This substantially strengthens the interpretability of the model.

Comment 5C.
Scaling analysis: Non-dimensionalize: show dependence on key ratios (e.g., motility vs adhesion).
Response:
We appreciate this suggestion. In the revised manuscript, we now frame the threshold interpretation more explicitly in terms of the competition between effective motility/protrusive forcing and

adhesive mechanical resistance, especially in the Results and Discussion. While we did not add a full formal nondimensional derivation, the new robustness analysis across adhesion strength directly addresses the reviewer's underlying point by showing how the threshold boundary shifts as the balance between motility and adhesion changes.

Comment 6.
Add explicit limitations: Stimulus is not mechanistically grounded, no modality-specific predictions, no biochemical signaling included, no experimental validation.
Response:
Implemented. We added a dedicated **Limitations** paragraph in the Discussion explicitly stating that:
(i) the control input is phenomenological,
(ii) the model does not include acoustic wave propagation or mechanotransduction,
(iii) the model does not generate modality-specific predictions,
(iv) biochemical signaling and matrix remodeling are not included, and
(v) the present work is validated only against published baseline scratch-assay data rather than new experiments.

Comment 7.
Quantify stochastic variability: run N ≥ 20–50 independent simulations … report mean trajectory, variance, confidence intervals …
Response:
We thank the reviewer for emphasizing this. We added a new Methods subsection, **"2.4 Stochastic simulation, uncertainty quantification, and numerical convergence,"** in which we specify that all reported wound-closure metrics were computed from **N = 30 independent simulations with distinct random seeds**. We now report ensemble mean trajectories and bootstrap 95% confidence intervals, and the updated figures and captions reflect this uncertainty-quantification framework.

Comment 8.
Report convergence checks: Does increasing number of cells / grid resolution change results? Report sensitivity to randomness …
Response:
We have addressed this point in the revised **Methods 2.4** section. There we now state that we repeated key simulations across different PDE grid resolutions and ABM initial agent counts and verified that the qualitative conclusions, including the threshold behavior, were robust. Because the current revision prioritizes the main biological and mechanistic results in the main text, these convergence checks are described in the Methods rather than shown as a separate main-text figure.

Comment 9.
Code release: Provide full repository … scripts to reproduce figures … seed control …
Response:
We appreciate this recommendation. We now explicitly document seed control and random-seed usage in the Methods. In addition, we will provide the simulation code, parameter files, and figure-generation scripts in a reproducibility package / repository accompanying the revised submission, as soon as the paper is published.

Comment 10.
These revisions will actually make the manuscript designable for a very high level impact … fit poor vs good wound-healer datasets …
Response:
We sincerely appreciate this forward-looking suggestion. We agree that extending the framework to fit wound-healing datasets from impaired versus healthy healing regimes would be a valuable next step and could make the threshold biologically more meaningful. We have not added this analysis in

the current revision because it would require a substantial expansion in scope and careful acquisition of appropriate time-series datasets. However, we have incorporated this idea into the discussion of future directions and believe it is an exciting avenue for subsequent work.

Reviewer 2
Comment 1.
Please verify that all links to the references point to the correct source.
Response:
Thank you. We manually checked all references and DOI links and verified that they point to the correct original sources. We also made minor formatting corrections to ensure consistency with the journal's reference style.

Comment 2.
The authors must disclose and acknowledge any assistance received in the preparation of this manuscript … All such contributions must be clearly stated in the Acknowledgments section.
Response:
We appreciate this comment and have added an **Acknowledgment** section to disclose editorial assistance and computational support used during manuscript preparation. The revised text clarifies that only minor language-editing/formatting assistance was used and that all scientific content, analysis, and interpretation were developed and verified by the authors.

Comment 3.
Include enough recent references along with foundational ones. Ensure references directly support your claims. Avoid "padding." Use credible sources …
Response:
Thank you. We reviewed the reference list carefully and revised it to ensure that it includes both **foundational references** (e.g., Fisher, Kolmogorov–Petrovsky–Piskunov, Sherratt & Murray, Maini et al.) and **more recent wound-healing / computational modeling references** directly relevant to the claims in the manuscript. We avoided adding unnecessary citations and retained only references that support the conceptual framing, modeling framework, wound-healing assay interpretation, or hybrid/agent-based methodology.

We again thank both reviewers for their constructive comments. Their suggestions significantly improved the conceptual framing, rigor, and transparency of the manuscript.

Thank you for trying to address my comments. However, my core comments have not been addressed.
1.Even after ostensible revisions, the paper still essentially represents a computational exploration of a model, not a validated model of wound healing. Even after revision, the study still follows this logic:
Choose a hybrid Fisher–KPP + ABM model
Select parameter values
Run simulations under different "control schedules"
Observe thresholds and dynamics
However, none of the key model components are grounded in experimental measurement. The only attempt at validation is the Fisher–KPP baseline calibration, which is weak because you state "representative literature control data from scratch-assay experiments" , but you do not But they do not specify the exact dataset, experimental conditions, measurement uncertainty, fitting procedure and parameter confidence intervals. Hence, you have not addressed my core concern; which is that the model parameters are not identifiable from real wound-healing experiments.

Note that agent-based models can generate almost any behavior depending on parameter choice. Without data constraints, these outcomes are not predictive, they are only illustrative simulations and not publishable.
You need to fit the model to REAL wound closure curves from the literature. Then estimate parameters and ask if the hybrid model improve fit vs Fisher–KPP alone?
Use experimentally measured values (again from literature) of cell migration speed, proliferation time, adhesion forces… and then check if the model reproduces closure dynamics.
Fit to control experiments, then predict perturbation experiments such as growth factor treatment, underlying pathophysiology (again from literature) and see if predictions match experimental data.
2.In my earlier review, I observed that " your threshold may be modeled by construction — not discovery". You have not addressed this. The threshold phenomenon you highlight may simply be a property of Fisher–KPP traveling-wave dynamics and/or force-balance between motility and adhesion. Hence, the result might be mathematically inevitable, not biologically discovered (which is why you need to fully address point 1).
3.The best practice of code release is to release it on submission as a DOI or Github link.
4.Convergence checks are not adequately addressed. You state that you performed grid-resolution tests. Where are the results? You do not present any tables, figures or quantitative comparison.
5.For parameter robustness, you have only varied one parameter (adhesion). Please perform robustness on D, ρ, drag and noise; as previously reviewed and communicated.
6.You did not address my scaling / nondimensionalization comment. Without scaling analysis, the result may simply reflect arbitrary parameter choices.With scaling, you can turn the simulation study into actual theoretical insight about wound closure dynamics. Scaling analysis often reveals which parameters matter and which do not. It reveals why thresholds matter. You can actually show something like ""Closure occurs when motility exceeds adhesion by a factor of ~2." That would be a general scientific insight. It would apply to different experiments, different cell types and different models.
Instead of running more simulations with different numbers, I would like you to explain the simple rule that actually controls wound healing in the model. Right now, your result is "Closure happens when amplitude and duration are high." However, those two parametes are arbitrary inputs. Scaling would show that "Closure happens when driving motility exceeds adhesive resistance by a factor of x" (or something along those lines). This turns it into a mechanistic explanation.
If - on the other hand - you were to find that the threshold is mathematically inevitable, or that the result does not depend strongly on the hybrid model and a simple Fisher–KPP + drag balance predicts it equally well, that would substantially reduce the novelty of your hybrid model to be published as a stand-alone independent manuscript.

Response to Reviewer

We sincerely thank the reviewer for the detailed and thoughtful critique. The reviewer raised important concerns regarding model validation, parameter identifiability, numerical convergence, and the mechanistic interpretation of the observed threshold behavior. We misunderstood the reviewer's initial comment at the first time, but we now fully managed the reviewer's concerns. In response to these comments, we substantially revised the manuscript and added several new analyses designed to address the reviewer's core concerns. These revisions include:

- explicit validation against literature wound-closure data
- quantitative comparison between Fisher–KPP and the hybrid PDE–ABM model
- numerical convergence analysis
- expanded parameter robustness analysis
- a nondimensional scaling analysis revealing the mechanistic condition governing wound closure

Below we respond to the reviewer's main points in detail.

1. Model validation using real wound closure data

Reviewer comment:
You need to fit the model to REAL wound closure curves from the literature.

Response:
We agree with the reviewer that a comparison with experimental wound closure data is necessary to establish a meaningful baseline validation of the model.
To address this concern, we introduced a new analysis in Section 3.1 of the revised manuscript using scratch-assay data reported in Johnston et al. (2015). Since the original study presents the wound closure curves graphically, the data were digitized from the published figures and converted into tabular form.
Both the Fisher–KPP continuum model and the hybrid PDE–ABM model were fitted to the extracted wound-closure trajectory. The resulting comparison is shown in Figure 1.
The hybrid model reproduces the observed wound-closure dynamics with slightly lower error (RMSE = 0.010) than the Fisher–KPP model (RMSE = 0.017). This result suggests that incorporating explicit cell–cell interaction mechanics improves agreement with experimentally reported closure dynamics.
This analysis provides a literature-based baseline validation of the hybrid modeling framework.

2. Hybrid model vs Fisher–KPP baseline
Reviewer comment:
Then estimate parameters and ask if the hybrid model improve fit vs Fisher–KPP alone?

Response:
Following the reviewer's suggestion, we explicitly compared the predictive performance of the hybrid PDE–ABM model against the classical Fisher–KPP continuum model using the same literature dataset.
As shown in Figure 1, both models reproduce the general sigmoidal wound-closure dynamics. However, the hybrid model achieves a lower RMSE relative to the digitized experimental data. This comparison demonstrates that the hybrid framework captures additional mechanistic effects—specifically cell–cell mechanical interactions—that are absent from the purely continuum Fisher–KPP model.

3. Numerical convergence analysis
Reviewer comment:
Convergence checks are not adequately addressed. Where are the results?

Response:
We missed the quantitative results about the following comments. In the revised manuscript we therefore introduced Section 3.5: Numerical Convergence, Parameter Robustness, and Scaling Analysis.
To verify that the results are not artifacts of discretization, we performed convergence tests for both the continuum and agent-based components of the model.
Specifically, we varied:

- the PDE grid spacing Δx between 5 μm and 20 μm
- the ABM resolution between 30 and 120 agents per side

For each configuration we measured the time required to reach 95% relative wound density.
The results are presented in Figure 7, which shows that the predicted closure time changes by less than 2% across grid resolutions and remains essentially unchanged across ABM resolutions.
These results confirm that the hybrid simulation results are numerically converged.

4. Parameter robustness analysis
Reviewer comment:

Please perform robustness on D, ρ, drag and noise.

Response:
Following the reviewer's recommendation, we expanded the sensitivity analysis to include variations in the following parameters:

- effective motility (D)
- proliferation rate (ρ)
- drag coefficient (γ)
- stochastic noise amplitude

For each parameter variation, we computed the threshold stimulus amplitude required to achieve wound closure within 48 h.
The results are summarized in Figure 8.
The closure threshold remains largely stable across variations in motility, proliferation rate, and noise amplitude, indicating that the threshold phenomenon is robust to these parameters. The threshold shifts only with the drag coefficient γ, which is physically expected because drag directly determines the mechanical resistance opposing migration.

5. Scaling analysis and mechanistic interpretation
Reviewer comment:
Without scaling analysis the result may simply reflect arbitrary parameter choices.

Response:
We agree with the reviewer that a nondimensional scaling analysis can provide a clearer mechanistic interpretation of the threshold behavior observed in the simulations.
To address this point, we introduced a dimensionless drive-to-resistance ratio, which represents the balance between migration-driving forces and mechanical resistance in the tissue.
Simulation outcomes across a wide range of parameter combinations were plotted as a function of this dimensionless parameter. As shown in Figure 9, closure outcomes collapse onto a single logistic transition curve when expressed in terms of Π.
This result demonstrates that wound closure occurs when the migration-driving forces exceed the mechanical resistance by approximately a critical factor. In other words, the threshold behavior observed in the simulations reflects a mechanistic balance between motility-driven migration and adhesive resistance, rather than an artifact of arbitrary parameter selection.

6. Code release
Reviewer comment:
The best practice of code release is to release it on submission as a DOI or Github link.

Response:
We archived the code and the dataset in
https://github.com/chengminsun09/Acoustic_Modelling_Wound_Closure.

Thank you for addressing my concerns. The manuscript is much improved. There are some deficiencies remaining.
1.You have (weakly) addressed my concern about model validation against real data. What remains problamatic is that you have only used one dataseet; which is digitized data from plots; which in turn introduces uncontrolled error.
-No uncertainty model for digitization noise
-No independent validation dataset (only fitting, no hold-out)

-RMSE difference is small and not statistically tested
Please correct these deficiencies.
2.Please use a second dataset and perform model validation on this second dataset also - one that does not require you to used digitized data (actual data can be extracted or is presented). See for example: 10.1186/s13742-015-0049-6
10.5524/100151
10.1016/j.csbj.2020.02.023
10.1186/s12860-021-00369-3
10.1186/s12918-015-0182-y
10.1016/j.ces.2018.01.004
10.1089/107632704323061834
10.1016/j.ces.2009.05.011
10.1016/j.jtbi.2014.01.018
10.1098/rsif.2013.0456
3.Hybrid vs Fisher–KPP comparison is missing model selection criteria (AIC/BIC/likelihood), missing statistical significance test, missing out of sample prediction.

_____

Response to Reviewer

We thank the reviewer for the careful follow-up assessment. We agree that our previous revision still relied too heavily on a single digitized dataset and did not separate in-sample fitting from independent validation clearly enough. In response, we substantially revised both the manuscript text and the validation analysis.

- The Johnston et al. (2015) curve is now treated only as an approximate calibration benchmark because it is derived from a digitized plot rather than from raw numeric data.
- We added a perturbation-based digitization sensitivity check showing that small extraction errors do not materially change the qualitative in-sample ranking for the Johnston benchmark.
- We added an independent validation using the public MDCK wound-healing time-lapse dataset of Zaritsky et al. (2015; dataset DOI: 10.5524/100118), which provides replicate-level OME-TIFF data and therefore does not require plot digitization.
- We introduced leave-one-replicate-out validation within each condition and now report out-of-sample RMSE and R^2, together with likelihood-based AIC/BIC for the training fits.
- We added a formal paired comparison of held-out RMSE using bootstrap confidence intervals and Wilcoxon signed-rank tests, and we revised our claims so that we no longer overstate hybrid-model superiority when the differences are modest.

Reviewer comment:

Thank you for addressing my concerns. The manuscript is much improved. There are some deficiencies remaining. You have (weakly) addressed my concern about model validation against real data. What remains problematic is that you have only used one dataset, which is digitized data from plots, which in turn introduces uncontrolled error. Please correct these deficiencies. Please use a second dataset and perform model validation on this second dataset also—one that does not require digitized data. Hybrid vs Fisher–KPP comparison is missing model selection criteria (AIC/BIC/likelihood), missing statistical significance test, and missing out-of-sample prediction.

Response:

We agree with the reviewer. In the revised manuscript, we restructured the validation section so that the digitized Johnston et al. (2015) dataset is no longer used as the primary basis for validation. Instead, it is presented only as a secondary, approximate calibration benchmark in Section 3.1 (Figure 1). For this benchmark, we now report not only RMSE but also AIC and BIC. The hybrid PDE–ABM model fits the digitized Johnston curve better than Fisher–KPP in-sample (RMSE 0.010 vs 0.017; AIC −75.80 vs −69.10; BIC −75.01 vs −68.70), but we now state explicitly that this result should be interpreted cautiously because the data were digitized from a published plot. To address the reviewer's concern about uncontrolled extraction error, we also added a perturbation-based sensitivity analysis for the digitized Johnston benchmark. Under small random perturbations of the

extracted trajectory, the hybrid model retained the lower RMSE in approximately 89% of realizations, indicating that the qualitative in-sample ranking is reasonably stable, although we continue to treat this benchmark as secondary evidence only.

To address the reviewer's main concern directly, we added an independent validation using the public MDCK time-lapse wound-healing dataset of Zaritsky et al. (2015; GigaDB DOI: 10.5524/100118). We analyzed five control replicates and five HGF/SF-treated replicates and extracted normalized wound-closure trajectories directly from the raw OME-TIFF image stacks. We then performed leave-one-replicate-out validation within each condition. These results are presented in new Section 3.2, new Figure 2, and new Table 1.

For the independent MDCK dataset, the hybrid model achieved slightly lower mean hold-out RMSE than Fisher–KPP in both conditions: control, 0.1129 vs 0.1156; HGF/SF, 0.1472 vs 0.1487. The hybrid model also yielded slightly higher mean hold-out R^2 (control 0.864 vs 0.857; HGF/SF 0.712 vs 0.704). However, the advantage was modest. Bootstrap confidence intervals for the RMSE difference included zero, and paired Wilcoxon tests were not significant (control $p = 0.4063$; HGF/SF $p = 0.0938$). We therefore revised the manuscript to say that the independent dataset supports the hybrid model as mechanistically richer and competitively predictive, but not decisively superior statistically under the present validation data.

These changes address the reviewer's concerns regarding dependence on a single digitized dataset, the lack of independent validation, the absence of out-of-sample prediction, and the need for likelihood-based model selection criteria and statistical comparison. We also updated the Discussion and Limitations sections so that the validation claims are fully aligned with the new results.

Thank you for addressing my comments. Accepted.