We thank the reviewers for the time and deliberation they put into their detailed comments. Our responses follow each in turn. If necessary, we are ready to expand upon any of the new content, which we tried to keep brief to avoid lengthening the paper.

First, we note that, prompted in part by the reviewers’ comments, we changed how we categorized each experiment by its task: Rather than the general “prediction”, we recoded these experiments to one of the other existing tasks or to “forecast” (i.e. predicting future events), and we changed the coding “decision support” to “recommendation”. The resulting tables and figures contain unhighlighted changes, but we have highlighted their captions as with Figure 1.  
  
Reviewer 1

Background  
  
The statement  "Localized models thus provide an individualized way to fit fully interpretable models to large population data" lacks clear supporting evidence. The authors should either provide substantial justification within the paper or revise the statement to reflect the actual findings more accurately.

We have amended the paragraph leading up to this sentence for clarity, and on reflection we feel that this specific sentence is unhelpfully abstract and so have removed it.  
  
Methods  
  
The authors state, "we focus on methodologies rather than conceptual approaches or evidence." However, in Section 4, they mention, "we focus here on the conceptual." This inconsistency should be clarified, and the manuscript should maintain a consistent focus throughout

Both uses of “conceptual” here have been replaced with clearer language that we believe prevents confusion.  
  
Results  
  
Why do the authors believe CBR is becoming more common? Although it is mentioned in the manuscript, a proper discussion on the underlying reasons is missing and should be included.

To be clear, we cannot say that CBR is becoming more common, but our limited evidence indicates that studies using similarity cohort–based models (renamed from “localized models”) have increased slightly in frequency over the years covered in our review. We have added some informed speculation to this point.  
  
The authors should provide a discussion on the types of data used and their influence on the performance and applicability of the discussed models. Understanding how different data types impact model behavior is crucial for evaluating the study's findings and generalizability.

While our sample is small and the models and experiments reported are highly heterogeneous, we have added some discussion and visualization of the relationships between performance and several aspects of each experiment.  
  
3.3 Application domains  
  
The authors did not include a neural network (NN) model in their analysis; however, towards the end of the paper, they mention having incorporated NN-based models. To ensure clarity and consistency, the authors should align their statements throughout the manuscript and provide a clear rationale for their model selection.

We apologize for the confusion—in our abbreviation scheme, “ANN” now refers to artificial neural networks while “NN” has been replaced with “kNN” to refer to nearest neighbor models.  
  
Conclusions  
  
The authors should provide a more in-depth and extended discussion on the identified needs, elaborating on their implications, potential challenges, and how their proposed approach addresses these gaps.

We have written a paragraph in the section describing our general framework that connects it back to the needs described in the Results.  
  
The term "localized models" can be ambiguous, as it may refer to models that focus only on a specific aspect of a disease, rely on features derived from a particular (or local) lab measurement, or, in the case of imaging, emphasize specific patches or kernels. To avoid potential misunderstandings, the authors should consider using a more precise term that accurately reflects their intended meaning.

Previous reviewers also raised this point, and these new points make it clearer still that the shorthand is confusing. We have replaced it throughout with “similarity cohort–based models”, abbreviated SCBM.  
  
Improving the user interface is not typically within the scope of academic research. It is important to distinguish the focus of research institutions from that of industry. Scientists engaged in research should not be expected to allocate time to developing user interfaces for individuals who are generally not involved in the research topic.  
  
The beginning of Section 4.5.2 should be rewritten to reflect this issue more clearly.

This point is well taken, and we have dropped mention of user interfaces and experience from this section.  
  
  
Reviewer 2

I appreciate the opportunity to review this paper. Unfortunately, I found it challenging to follow. I encourage the authors to improve their work and resubmit it. Below, I provide detailed comments to help enhance the paper.  
  
1. The abstract suggests that the literature search included papers published between 2021 and 2024. However, it later mentions including papers from 1997 to 2022. Since the authors initially stated that they limited their search to papers published after 2021, this raises concerns that relevant studies might have been omitted.

We have rewritten this part of the abstract in what we hope is clearer language.  
  
2. The abstract should separate the objective from the background.

Thank you for catching this; done.  
  
3. The methodology section of the abstract needs to specify the eligibility criteria for the included literature.

Without enumerating the criteria, we have summarized them, as we hope is appropriate for the abstract.  
  
4. The results section of the abstract should be written more scientifically. Phrases such as "almost all" and "most" are vague and not informative. I encourage the authors to proofread their work carefully. When revising the document, they should assume that readers may wish to replicate the study, which is why the paper must be written with scientific rigour.

This point is well-taken. We have quantified each of these statements.  
  
5. The introduction makes strong claims without proper citations, which may affect the credibility of the arguments. The following statements, for example, need supporting references:  
  
"which appears to have been introduced several times independently under different names, and in slightly different forms."  
"AI models often process hundreds or more variables in opaque ways to yield predictions that depend on not only their values but their associations and interplay with each other."  
"The loss of straightforward interpretations of these models has limited their practical uptake and motivated the construction of explanatory statistics for opaque models as well as the development of more interpretable complex models."  
"we call localized modeling."

With the exception of the last (“localized modeling” has been renamed “similarity cohort–based modeling”), these and several other assertions have been (sometimes rephrased and) referenced in the revision.  
  
6. The authors should ensure clarity and precision in their writing. For instance, the acronym EHR is used in the Background section without being defined.

Thank you for catching this. We believe that all initialisms are now introduced at their first use and included in the list at the end (some having changed with the revision).  
  
7. The discussion on AI in the Background section does not mention Explainable AI (XAI), which is an important aspect of the field.

This is a very good point. We have added a paragraph distinguishing interpretable and explainable ML with our reasons for focusing this review on a method we understand to fall into the former type.  
  
8. The Related Work section lacks references and should be supplemented with relevant citations.

9. The authors should compare their work to existing review papers and explain why their review is necessary.

We have added several new references to this section and some additional sentences expressing the gap in the literature that we aim to fill. The specific references added are other review and survey articles that helped us articulate our goals and carve out our niche. We don’t suggest that it is exhaustive but believe that it will allow readers a clear entrance into any of these closely related topics.  
  
10. The structure of the paper needs revision. For example, the selection criteria are currently placed in the Objectives section, which is unexpected.

We have moved the eligibility criteria to the Methods section and removed the recap of them there previously. Driven by another comment, we also renamed the Conclusions section to Discussion, renamed the Closing subsection to Conclusions, and shortened the Discussion and eliminated several subheadings. We did not see any other major structural problems with the document, but we will be glad to make any more that are pointed out.  
  
11. The Methods section refers to another section rather than presenting key methodological details clearly. The authors should reconsider how they organize and present their work.

This was due to a cross-referencing error in LaTeX and has been fixed. In the interest of length we did not include all details of the search, screen, review, and synthesis in the manuscript but left some more technical details in the Appendix. We will be glad to relocate some of these details if the reviewers think it appropriate.  
  
12. The choice of search engines should be justified. Why were the selected databases chosen? Why was IEEE not included?

We have commented on the reasons for our original choices. IEEE was not included because the authors were not familiar with it. We are ready to adapt the search and update the results and analysis if the reviewers believe that this database is essential, but we would require an estimated additional 14 days to complete it.  
  
My overall observation is that this paper would benefit from a reduction in length. The authors should also focus on interpreting the reviewed literature rather than merely summarizing it. Additionally, the Conclusion section is not informative. Furthermore, there are several sections following the Conclusion, which is unexpected.

We have reduced the length of the manuscript by excising text we find less essential or moving it to the Appendix. Largely in response to the reviewers’ comments, we have also added several paragraphs that do more to interpret our results. And, as noted above, we have renamed and reorganized the final sections, which now comprise Discussion and Conclusion.  
  
I hope these comments help the authors improve their manuscript.

Indeed, we believe they have, and thank you.