Many thanks to all the reviewers, but I can see that we've split the field. There seems to be consensus on the problem domain (ML implementation), but you've raised challenges as to whether our solution is *research*, whether we have appropriately addressed concerns with *privacy*, our lack of *technical detail*, and that the abstract is *off topic with no ML or healthcare application.*

Themes from more than one reviewer:
This is not research: We agree. But given that this year's symposium highlights *'real-world robustness and generalization'*, and observes that *'models ... underperform dramatically, unexpectedly, and inequitably when integrated into care processes'*, we argue that a discussion of a proposed solution that brings the care process into the development loop is an important contribution. We would also respectfully suggest that just as there is now consensus that software development is a valid academic contribution (`https://www.software.ac.uk/about/manifesto`) then the same standard should apply to platform development. Our contribution is generalisable, and open source (shared values from scientific research), and we are bringing it to the conference to seek the views of the wider research community.
Off topic/No healthcare or ML application: See comments above especially that we see the ML4H researcher as our primary user, and the design aims to improve the efficacy/impact/efficiency of their work.
Privacy: We use the '5 safes' framework and *'code to data'* delivers a *safe setting* (data remains in the hospital), and *safer data* (because we control the triangulation risk with external data sets). Both EMAP databases and the development environment are hosted *within* the hospital firewall (private cloud). Users are under the same governance process as all hospital staff etc. There is no *external* access. Users do not access raw IDS messages. Authorised developers may work with access to PII (Personally identifiable information) for application deployment but research teams work with a de-identified view of the UDS. Queries and database access are logged.

Specifics including technical detail, future work and formatting
1. EMAP data grows at 30GB/month/1000 beds (i.e. notes/structured data) equivalent to 5 million HL7 messages with a message latency of 3.92 (IQR $0.9 - 6.7$) seconds. Imaging etc. stored as links to clinical image repository. **2.** IDS has fully redundant replicates and the UDS is designed to be rebuilt at will from the IDS. **3.** Ongoing integration with local `https://cogstack.org` messages with free text sent to Cogstack and the NLP concept returned and stored. **4.** We will remove Figure 1, 3 & 4 and use the space to add technical detail (parenthetically), system performance (paragraph), and exemplar model performance (paragraph): final document (4 pages) and Appendix (1 page).

Response to individual reviewers (Thank you all for the considered critiques.)
Review A: Thank you. We know only of <5 hospitals that have an online development process (e.g. Duke/Yale/Columbia). Some of these are proprietary. Our EHRS vendor (Epic) tell us we are the only such customer to have live data. Prominent published models (Hyland, Nature Med, 2020; Tomasev, Nature, 2019) were developed offline and have *not* been not been deployed. And those deployed (Sendak, App Clin Inform, 2017; Razavian NPJ Digit Med, 2020) were expensive or unsuccessful. We agree there are many forms of EHR but nearly all use HL7 integration engines, and work with HL7v2.3+ so this is widely generalisable. We acknowledge the utility of NoSQL for healthcare, but we chose SQL because it means that the system is usable by the hospital business team as well as researchers thus augmenting its utility/community. We apologise for the lack of detail and would commit to improving this.
Review B: Thank you for recognising the issues of not disturbing operational systems. You asked about other reported solutions: these include Sendak et al. (JMIR Med Inform, 2020), and (McPadden et al. (J Med Internet Res, 2019). We prioritised space (platform over model) but can supplement in an appendix.
Review C: Thank you. We agree that non-technical issues are major barriers to deployment, but we'd argue that its exactly for the issues of safety that EMAP is valuable. We isolate the operational system, and we enable live shadow deployment to generate safety data, examine model stability, and iterate over the target implementation. Small details (e.g. sending emails rather than hosting a dashboard; shortened forecast horizons) improved the uptake of our local models, and these are more effectively managed during development. We agree that data standardisation is key (e.g. FHIR), and are working on an OMOP implementation.
Review D: Thank you for noticing our emphasis on 'developer ergonomics (e.g. the hosted notebooks and SQL-as-an-API abstraction)'. EMAP is open source with the specific intention of building a community of use. Documentation is a priority but not yet to the standard we intend.
Review E: Thank you. We chose to focus on the system rather than the case study because with limited space this is the novel contribution. We apologise for not explaining the safety management more clearly. EMAP is a system for rapid prototyping. Live models are deployed but (1) as web apps with user access control limited to clinicians working with the research team, or (2) model outputs returned into the EHR labelled as research not clinical 'grade' data with separate access control (work-in-progress).