MSR79

Deploy-AI: A Framework and Future Checklist for Facilitating Deployment of Clinical Prediction Models

JOULES A[1], BRUSINI I[1], ALI R[1], JANI M[2], BROWN B[2], PEEK N[3], RIGG J[1], MCBETH J[4], DIXON W[2]

[1]IQVIA, London, UK, [2]University of Manchester, Manchester, UK, [3]University of Cambridge, Cambridge, UK, [4]University of Southampton, UK

Motivation

The current literature *lacks guidance* on how to design algorithms with downstream clinical deployment considerations in mind.

Our Solution

We propose the compilation of **Deploy-AI**: a novel checklist of deployment considerations that can guide algorithm design and ensure that clinical prediction models are *usable* and *fit-for-purpose*.

Case Studies 1 and 2

An algorithm trained on US ambulatory electronic health records (EHR) to find undiagnosed patients with hepatitis C virus (HCV)[1] needs to be deployed on Australian EHRs as a clinical decision support tool (CDST) for HCV screening[2].

Another algorithm, trained on linked primary and secondary care data in Salford, UK, will find axial spondylarthritis (axSpA) cases, and be deployed in the UK Greater Manchester Care Record to aid early diagnosis.

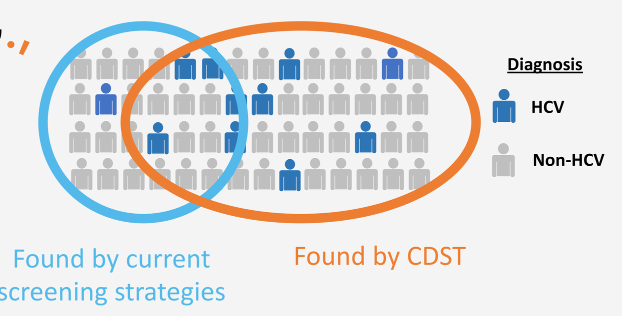
Case Study 1

In 2020, 117,810 Australians had chronic HCV[3].

The World Health Organization aims to eliminate HCV by 2030 through proactive screening and treatment[4].

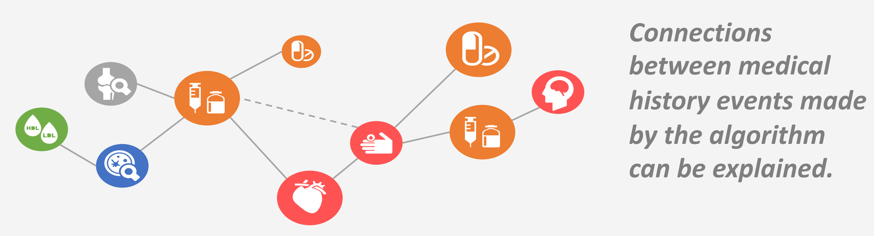
The algorithm should show a meaningful clinical benefit over current screening strategies, by identifying patients in subgroups that aren't currently targeted or known to be at-risk.

E.G.,



Patient need for HCV screening is estimated using their clinical history

The algorithm's results and workings should be explained to the CDST user via a dashboard:



The dashboard should provide relevant medical information and patient contact details, updated weekly.

The algorithm can be an XGBoost model with SHAP-based feature importance to highlight key clinical features influencing patient prioritization.

The Australian government has guidelines for SAMD and exempt software[5,6]

The algorithm should not be "black box" and the CDST should include model explainability metrics. The intervention decision should be made by the healthcare professional.

EHR data is hosted in a Microsoft Azure SQL database inaccessible to model developers

Codebase should be containerized following best engineering practices for easy installation and configuration.

Codebase should be connectable to SQL database, adapted to the Australian EHR data schema, and set up to run efficiently (e.g., *an off-peak schedule*).

A virtual machine will host the algorithm.

Deploy-AI

Clinical considerations

Is the CDST appropriate for clinical deployment?

- ❑ Why do we need the CDST and how will it improve the current standard of care?
- ❑ What is the benefit to patients and healthcare providers?
- ❑ Who will use the CDST and where will it sit in current clinical practice?
- ❑ What information will the end user need to make an informed intervention?
- ❑ Is there clinical capacity for this CDST?

Regulatory considerations

Is the CDST safe for clinical deployment?

- ❑ Algorithmic bias and fairness
- ❑ Data privacy frameworks
- ❑ Software as a Medical Device (SaMD) considerations
- ❑ AI accountability principles
- ❑ Information governance

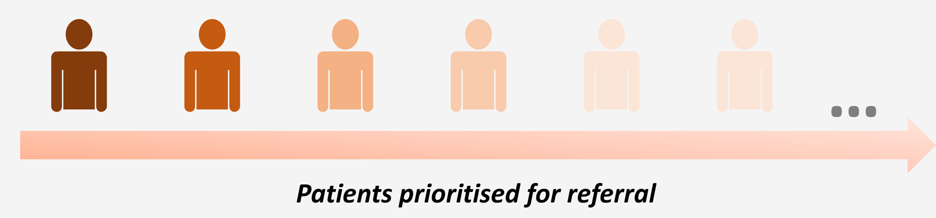
Technical considerations

Is the CDST feasible for clinical deployment?

- ❑ In which infrastructure will the CDST be deployed?
- ❑ What are the minimum technical requirements to deploy this algorithm?
- ❑ Which data will feed the CDST? Is it expected to differ from the training data, and how can this be addressed?

Case Study 2

A population-health dashboard prioritises patients for general practitioners to assess suitability for axSpA referral



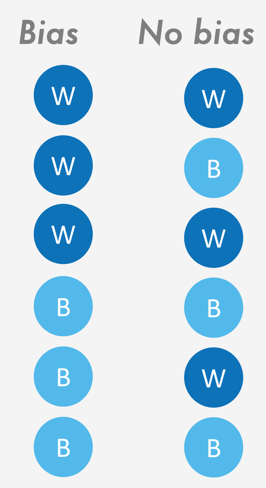
Patients prioritised for referral

Medical information for high-priority patients is shared and general practitioners decide on referrals to specialist centers.

The algorithm should:

- Process patient data on any chosen date;
- Be tested by a retrospective study that mirrors the deployment scenario, where training data temporally precedes the test dataset;
- Be optimised to achieve high precision on the highest prioritised patients.

AI systems should be inclusive and accessible, and should not result in unfair discrimination against individuals or groups[7]



Developers should:

Identify and report bias in the data

- Identify protected characteristics (e.g., gender, ethnicity, deprivation) in the EHR population.
- Analyse subgroup representation and disease incidence in the training dataset, to detect underrepresentation or underdiagnosis bias. *E.g., Salford's predominantly White population with higher deprivation levels may lead to underrepresentation of other ethnicities and deprivation levels, affecting algorithm performance in these groups.*

Identify and report bias in the algorithm

Analyse algorithm accuracy for each subgroup to detect discriminatory bias.

The structure of the Greater Manchester Care Record is known, but its content is inaccessible to developers

To ensure external generalisability and address differences between training and deployment datasets, adaptations are needed, e.g.:

- Mapping of clinical coding systems.
- Comparing variable distributions, if available.
- Exclusion of algorithm variables if poorly-captured in deployment data, and awareness of the impact on performance.
- Considering how differences in coding and clinical reporting practices may reduce algorithm precision. *E.g., local formularies may cause differences in medication prescriptions.*

Deploy-AI guides developers in designing algorithms that meet downstream deployment requirements and complements frameworks such as TRIPOD+AI, which focuses on reporting.
We aim to publish a finalised comprehensive checklist in future.

1.Rigg, J., Doyle, O., McDonogh, N., Leavitt, N., Ali, R., Son, A. and Kreter, B., 2023. Finding undiagnosed patients with hepatitis C virus: an application of machine learning to US ambulatory electronic medical records. *BMJ Health Care Inform*, 30(1), p.e100651. doi: 10.1136/bmjhci-2022-100651.
2.APASL, 2024. Abstract Book (Tentative) APASL 2024 Kyoto. Available at: https://www.apasl.info/history-meetings/apasl-2024-kyoto/docs/info/Abstract_Book_(Tentative)_APASL_2024_Kyoto_20240302.pdf [Accessed 6 November 2024].
3.Kirby Institute, 2020. Monitoring hepatitis C treatment uptake in Australia (Issue 12). The Kirby Institute UNSW. Available at: https://kirby.unsw.edu.au/sites/default/files/kirby/report/Monitoring-hep-C-treatment-uptake-in-Australia_Iss12-JUL22.pdf [Accessed 6 November 2024].
4.World Health Organization, 2024. Elimination of hepatitis by 2030. Available at: https://www.who.int/health-topics/hepatitis/elimination-of-hepatitis-by-2030#tab=tab_1 [Accessed 6 November 2024].
5.Therapeutic Goods Administration, 2024. Artificial intelligence (AI) and medical device software. Available at: https://www.tga.gov.au/how-we-regulate/manufacturing/manufacture-medical-device/manufacture-specific-types-medical-devices/artificial-intelligence-ai-and-medical-device-software [Accessed 6 November 2024].
6.Therapeutic Goods Administration, 2024. Clinical decision support software. Available at: https://www.tga.gov.au/sites/default/files/clinical-decision-support-software.pdf [Accessed 6 November 2024].
7.UK Government, 2024. Algorithmic Transparency Recording Standard Hub. Available at: https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub#:~:text=The Algorithmic Transparency Recording Standard,how algorithmic tools support decisions [Accessed 6 November 2024].

Contact information:
rehan.ali@iqvia.com