

Using predictive risk modelling to identify patients with hidden health needs in an Aboriginal and Torres Strait Islander health service

Gayani Tennakoon,
Rhema Vaithianathan,
Samantha L Pope, Zoe E Shiels,
Danielle C Butler, Lyle Turner

Background and objective

In partnership with an Aboriginal and Torres Strait Islander community-controlled health service, we explored the use of a machine learning tool to identify high-needs patients for whom services are harder to reach and, hence, who do not engage with primary care.

Methods

Using deidentified electronic health record data, two predictive risk models (PRMs) were developed to identify patients who were: (1) unlikely to have health checks as an indicator of not engaging with care; and (2) likely to rate their wellbeing as poor, as a measure of high needs.

Results

According to the standard metrics, the PRMs were good at predicting health checks but showed low reliability for detecting poor wellbeing.

Discussion

Results and feedback from clinicians were encouraging. With additional refinement, informed by clinic staff feedback, a deployable model should be feasible.

PRIMARY HEALTHCARE plays a central role in serving patients with complex care needs and providing person-centred holistic care to maintain and improve health and wellbeing. However, in primary care settings, for many reasons, patients who most need prevention services are often those least able to engage with services.^{1,2} The challenge is identifying patients with high but hidden healthcare needs.

Predictive risk models (PRMs) are increasingly used by health agencies, including in the Australian primary healthcare setting,^{3,4} to improve service provision to patients with high, often hidden, complex health needs, who might benefit from targeted proactive care.³⁻⁸ Existing predictive models developed in Australia tend to flag patients at risk of needing services for a narrow range of physical health outcomes so as to optimise the use of limited healthcare resources or to reduce preventable hospitalisations. In addition, studies show considerable bias in prediction models that focus on future resource use as the outcome.⁹ Although not directly examined with Australian data, this most likely applies to the Australian context as well. When health systems are incapable of addressing the needs of a particular population subgroup, those who do not use services might have higher needs than those who do.⁹ Moreover, in Australia, data used for developing these tools have been based on the general primary care population and might not be representative

of the Aboriginal and Torres Strait Islander population. The gap that we address in this paper is to use data from an Aboriginal and Torres Strait Islander community-controlled health service (ACCHS), and hence expect service use to better reflect actual need; then, among this group, we aim to identify those we think have high needs but are still not accessing services.

The Institute for Urban Indigenous Health (IUIH) is a community-controlled health service that leads the planning, development and delivery of healthcare, family wellbeing and social support services to the Aboriginal and Torres Strait Islander population of southeast Queensland. In collaboration with IUIH, prototype PRMs were developed to identify patients at risk of low engagement but with high health needs who could benefit from additional supportive and preventative care.

Methods

Study concept and design

This study was completed as part of a larger project implementing a redesigned model of holistic primary healthcare at the participating service.¹⁰ A working group oversees this project and includes Aboriginal and Torres Strait Islander and non-Indigenous researchers, clinicians, managers and community liaison officers. This group provides cultural and technical oversight of the project and related subprojects, including with respect to data

sovereignty, ensuring that what is measured is meaningful, culturally and clinically. This specific study was a collaboration with clinicians and service providers, who are also members of the community and were directly involved in the design and interpretation of the research.

A PRM uses a set of rules to summarise learned correlations. PRM tools are ‘trained’ using historical data to predict outcomes within a future time frame. During the training phase, the model learns the correlations between the features (or predictors) of the patients and the training outcomes. The resultant model can then be used to predict the corresponding outcomes (that the model was trained on) for the patients.

We used routinely collected deidentified data extracted from electronic primary healthcare records of the participating IUIH clinic. Participants included patients who had at least three occasions of service with the clinic in the two years prior to 1 March 2018. Extracted patient information data included demographics, previous appointments, medical history, medications prescribed, risk factors (eg alcohol consumption), and vegetable and fruit consumption. In this study, the period up to 1 March 2018 was considered as historical and the period from 1 March 2018 to 1 March 2020 as the future time frame where outcomes were observed.

As summarised in Table 1, we derived 377 predictor variables including all available data fields in historical electronic health records.

Outcomes

As outlined in the introduction, patients who most need prevention services are often those least able to engage with services.^{1,2} In this study, we aimed to identify patients who have a high risk of becoming disengaged but have features similar to those patients who are engaged and have high needs.

For the first part, we developed a tool to predict those patients at high risk of not coming for their next health check in a timely manner, as a proxy for disengagement. The clinic regularly invites and encourages all active patients to attend an annual health check involving a comprehensive assessment of the patient’s physical, psychological and social wellbeing.¹¹ However, many patients do not attend these health checks but might later present as acute patients.

Rather than restricting attention to a narrow rehospitalisation risk to identify those with high needs and at risk of preventable harms, we included a measure based on a patient’s own assessment of their health and wellbeing: self-reported poor overall health within two years. The specific question was ‘How strong (deadly) do you feel in your life at the moment?’, with responses ranging

from 0 (very poor) to 10 (excellent). Table 2 summarises the training outcomes of the PRMs, the size of the data sample and the prevalence rate, showing the percentage of patients with that outcome.

We did not have available outcomes for all patients in the study sample due to high levels of missing information in the medical records for the self-reported health measure (Table 2). Therefore, when training that model, only those patients for whom a self-reported health measure was collected were included.

Modelling

Model training

We trained two PRMs targeting the outcomes listed in Table 2, using the 377 predictor variables derived from historical data, as summarised in Table 1.

First, we partitioned the research dataset into a training set containing 75% of the data and a test set containing 25% of the data. Then we trained the PRMs using the Least Absolute Shrinkage and Selection Operator (LASSO) regularised logistic regression algorithm. The logistic regression method models the likelihood of a certain categorical outcome (0/1 in this case) as a function of a set of predictor variables. The LASSO regularised logistic regression method ensures an accurate and simple model by selecting significant predictor variables.¹²

Table 1. Predictor variables used for training predictive risk models

Data source	Predictor variables (sample)	Total no. predictor variables used for training the model
Demographics	Gender, age, Indigenous status, demographics missing	4
Appointments	Days since last appointment, count of previous appointments, ^A appointment history missing	6
Medical history	Count of medical history events, ^A count of different medical conditions, ^A medical history missing	189
Medication history	Count of scripted medicine for different conditions, ^A medication history missing	157
Alcohol consumption	Average drinks per day, ^B drinking frequency, ^B past frequency, ^B heavy drinking frequency, ^B alcohol history missing	9
Fruit and vegetable consumption	Fruit consumption per day, ^C vegetable consumption per day, ^C days to last fruit/vegetable consumption, fruit and vegetable consumption is missing	12

^APredictors are coded across the following time periods: within the last 90, 180 and 360 days, and ever.

^BPredictors are coded as both the most recent record and highest recorded value in history.

^COriginal values are categorical, which are transformed into a set of features such that each categorical value becomes a new feature with a 0/1 response.

The trained LASSO model consists of a set of weights related to each feature that can be used to calculate the probability of the targeted outcome. The LASSO models were trained using the R package named ‘glmnet’ (developed by Jerome Friedman, Trevor Hastie and Rob Tibshirani, Stanford University, Stanford, California, USA).¹³ Alternative modelling approaches of random forest¹⁴ and LightGBM¹⁵ were also tried, but there were minimal differences in terms of accuracy.

Model testing

We use the hold-out test dataset when evaluating the model’s performance. The LASSO model is applied to predict the probability of the targeted outcome. The predicted probability is a value between 0 and 1, whereas in this case, the observed outcome is a binary value of 0 or 1. To evaluate the accuracy of the model prediction, a cut-off threshold is considered such that the probabilities higher than the threshold are considered as 1 and the others are considered as 0. In addition to using the standard threshold of 0.5, we stratified the risk probabilities into 10 equally distributed risk scores where the top 10% of high-risk patients are assigned a risk score of 10.

Machine learning models are generally tested for accuracy using a set of standard

metrics, such as the area under receiver operator characteristics curve (AUC), the true positive rate (TPR) and the positive predictive value (PPV). The AUC measures a model’s ability to differentiate between patients who observed and did not observe the targeted outcome (ie positive and negative classes). A PRM with an AUC of 0.50 is no more accurate at identifying a patient at risk of an adverse outcome than a random guess. A PRM with an AUC of 1.0 can perfectly classify patients at risk of an outcome. The TPR tells us the proportion of patients flagged by the model out of those who experienced the outcome, which indicates the sensitivity of the PRM. Conversely, the PPV tells us the proportion of patients flagged by the PRM who go on to experience the outcome. TPR and PPV values make sense when presented with the cut-off thresholds.

Ethics approval for this project was obtained from The University of Queensland Human Research Ethics Committee (2021/HE002009).

Results

Data from 2258 regular patients as of 1 March 2018 were available for analysis. The average age of this population was

28 years, with 60.8% aged >18 years, 44% being male and 88% being Aboriginal and Torres Strait Islanders. Table 3 presents the performance matrices for the two PRMs built with LASSO. The model predicting ‘Not getting a health check’ shows a fair overall predictive power.

The models were trained using the full data sample available with respect to the outcome. Because the patient population of interest is those with an Aboriginal and Torres Strait Islander status of ‘Yes’ in their health record and aged ≥18 years, the remainder of this paper reports the models’ performance for that subpopulation.

Figure 1 plots the prevalence rate of not getting a health check against the PRM. Of patients assigned a risk score of 10 (ie highest risk), 78% did not have a health check, whereas 9% of patients with a risk score of 1 (ie lowest risk) did not have a health check. These results indicate that the PRM has a high precision when stratifying patients according to the risk of disengagement. The top 10% of high-risk patients predicted by the model accounted for 22% of patients who had not had a health check in the follow-up period 1 March 2018 – 1 March 2020, whereas for the top 50% of high-risk predictions, this value increased to 77%. This shows that the

Table 2. Details about outcomes used for building predictive risk models

Outcome (all within 24 months)	Definition	Sample size for whom outcomes were observed ^A	Prevalence rate ^B (%)
Not having a health check within 24 months	No patient record of having received a routine health check	2,258	40.26
Reporting a poor self-reported health score (<5)	Self-reported health measures are typically collected during the health checks. Self-reported health is coded on a scale of 1 (very poor) to 10 (excellent). Patients were coded as having poor self-reported health if they reported an average <5 (which was the bottom 25th percentile reported)	1,009	21.90

^AThis is the total number of patients where the outcome was observed. This includes both patients where the outcome was observed to be true and false.

^BThe prevalence rate is the percentage of those patients where the outcome was observed to be true.

Table 3. Performance metrics of Least Absolute Shrinkage and Selection Operator models

Outcome	AUC	TPR (threshold = 0.5)	PPV (threshold = 0.5)
Not getting a health check	0.77	0.67	0.63
Reporting a poor average self-reported health score	0.65	0.46	0.35

AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; TPR, true positive rate.

PRM has greater sensitivity for stratifying patients who are disengaging from services to higher-risk groups. For example, if we select the patients with a risk score of 5 or higher, it includes the 77% of patients who did not come back for health checks.

Based on the standard performance metrics used, the PRM is efficient at predicting patients who are at risk of future disengagement.

Although not having regular health checks (ie a person who might be unseen by

clinicians) is potentially a sign of high need, the second component is to identify those at heightened risk of preventable harms.

Figure 2 plots the prevalence rate of self-reported poor health against the PRM score. Of the patients assigned a risk score of 10 (ie highest risk), 63% self-reported poor health, whereas 44% of patients with a risk score of 1 (ie lowest risk) also self-reported poor health. These results indicate that the PRM has lower precision and was not reliable

in stratifying patients according to the risk of self-reported poor health.

Discussion

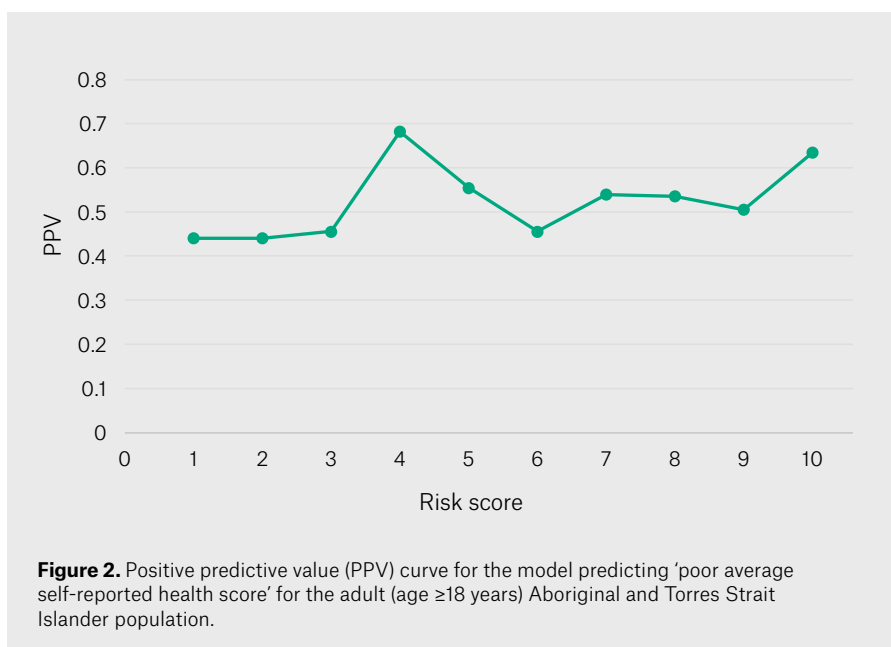
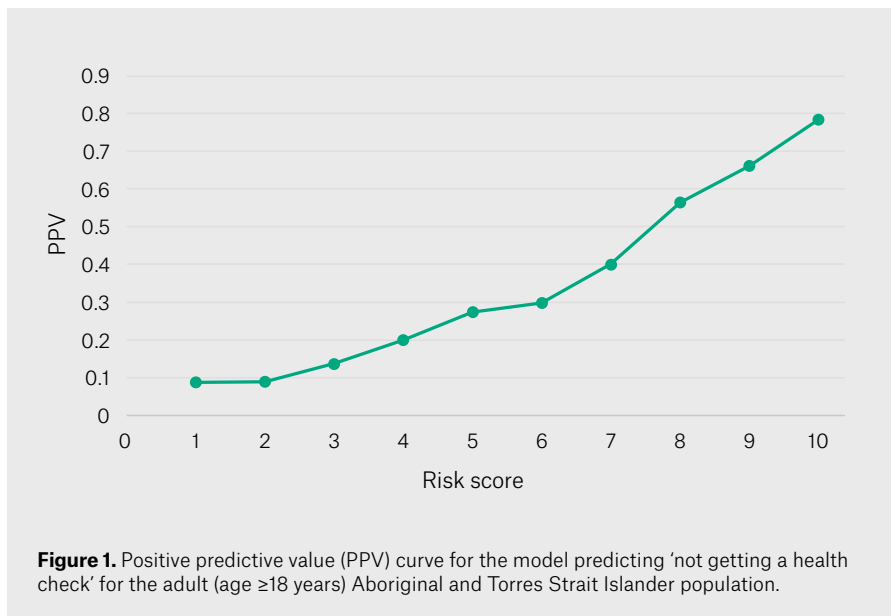
The objective of this research was to develop PRMs to identify patients who could benefit from proactive prevention in an urban ACCHS. According to the standard AUC, TPR and PPV metrics, the PRM performed well in identifying those who were less likely to engage with services. It was unreliable in detecting those who self-rated their wellbeing as poor because it was unable to discriminate between patients who were at high or low risk of rating their health as poor.

Strengths and limitations

To the best of the authors' knowledge, this is the first study to use routinely collected primary care data from an urban ACCHS to develop a PRM. The study established that these data are suitable for producing a reliable model. A strength of the study is that it uses a holistic measure of health, based on the patient's own assessment of their health and wellbeing. However, the completeness of this measure across the sample was lower than expected. When the data sample gets smaller, models have fewer examples to learn from and will not generalise well to new data. We used 'not having a regular health check' as a proxy for disengagement because annual health checks are a fundamental component of the IUIH system of care for all patients. However, there might be patients who are generally well and hence do not seek out this service. We were unable to distinguish this group from the main sample.

The use of data-driven risk-prediction tools in practice is challenging and might not always be fit for purpose. Bringing the technical and clinical aspects together is necessary to refine risk-prediction tools for clinical application and relevance. In this initial exploration, tools were developed in collaboration with providers, managers and clinicians, which has involved ongoing feedback and review to further refine and ensure utility.

There is also a risk of feature drift and degradation of the effectiveness of PRM tools over time. We plan to address this by undertaking quality assurance to determine whether refinement of the tool is required.



However, once tools are deployed, a lack of experimental data is always a challenge that needs to be overcome.

Further, this study was undertaken at a single study site; as such, the findings might not be generalisable to other settings.

Implications and next steps

In preliminary discussions and case reviews, clinicians were generally positive about the PRMs' potential but wanted to refocus the modelling efforts; for example, to identify patients who tend towards intermittent acute episodes of care, who are stabilised over a period and then are not seen again until they re-present in an unstable condition. Feedback from clinicians indicated that often these peaks in care need reflect social, emotional and mental health concerns, rather than physical wellbeing. With further refinement of the current model, such patients should be identifiable and, through additional iterative development with clinicians and clinic staff, a deployable model should be feasible.

Proactive prevention requires engagement with at-risk patients and their families when their acuity levels are low. Holistic comprehensive services like the ACCHS are well placed to respond in this way but might require actively seeking out and working with patients and their families who have not yet engaged with the service. In that way, services can work from a strengths-based perspective and there is more time to address social care challenges and build stronger engagement with providers across the board. These stronger connections then stand in good stead should future health crises arise, allowing health services to be more effective in responding to patient needs and priorities.

Authors

Gayani Tennakoon PhD, Postdoctoral Research Fellow, Institute for Social Science Research, University of Queensland, Brisbane, Qld; Research Consultant, Centre for Social Data Analytics, Auckland University of Technology, Auckland, New Zealand

Rhema Vaithianathan PhD, Professor of Health Economics and Director of Centre for Social Data Analytics, Centre for Social Data Analytics, Auckland University of Technology, Auckland, New Zealand; Professor of Social Data Analytics, Institute for Social Science Research, University of Queensland, Brisbane, Qld

Samantha L Pope, Kuungkari and Juru, Practice Manager, Institute for Urban Indigenous Health (UIIH), Windsor, Qld

Zoe E Shiels, Wakka Wakka, Cert 3 Primary Health Care, Aboriginal Health Workers, UIIH, Windsor, Qld
Danielle C Butler MBBS, MPH, FRACGP, PhD, General Practitioner and Health Services Researcher, UIIH, Windsor, Qld; Research Fellow, Australian National University, Canberra, ACT

Lyle Turner PhD, Manager Data and Research, UIIH, Windsor, Qld; Honorary Senior Fellow, University of Queensland Poche Centre for Indigenous Health, Toowong, Qld

Competing interests: None.

Funding: This research was supported by the University of Queensland Poche Centre for Indigenous Health through their Research Collaboration Seeding Grant and the University of Queensland AI Collaboratory funding through their Research Support Package.

Provenance and peer review: Not commissioned, externally peer reviewed.

Correspondence to:

rhema.vaithianathan@aut.ac.nz

Acknowledgements

The authors acknowledge the ongoing sovereignty of Aboriginal and Torres Strait Islander Peoples across our lands, and pay their respect to Elders, past and present. This research was conducted in partnership with the Institute for Urban Indigenous Health (UIIH). The authors acknowledge the UIIH System of Care Version 2 (ISoC2) Working Group and Moreton Aboriginal and Torres Strait Islander community-controlled health service. The authors would also like to thank Dr Tom Hilton for his insight and advice on the conceptualisation and interpretation of this study.

References

1. Rolewicz L, Keeble E, Paddison C, Scobie S. Are the needs of people with multiple long-term conditions being met? Evidence from the 2018 General Practice Patient Survey. *BMJ Open* 2020;10(11):e041569. doi: 10.1136/bmjopen-2020-041569.
2. Marmot M. An inverse care law for our time. *BMJ* 2018;362:k3216. doi: 10.1136/bmj.k3216.
3. Khanna S, Rolls DA, Boyle J, et al. A risk stratification tool for hospitalisation in Australia using primary care data. *Sci Rep* 2019;9(1):5011. doi: 10.1038/s41598-019-41383-y.
4. Pearce C, McLeod A, Rinehart N, et al. Polar diversion: Using general practice data to calculate risk of emergency department presentation at the time of consultation. *Appl Clin Inform* 2019;10(1):151-57. doi: 10.1055/s-0039-1678608.
5. Lewis G, Kirkham H, Duncan I, Vaithianathan R. How health systems could avert 'triple fail' events that are harmful, are costly, and result in poor patient satisfaction. *Health Aff (Millwood)* 2013;32(4):669-76. doi: 10.1377/hlthaff.2012.1350.
6. Curry N, Billings J, Darin B, Dixon J, Williams M, Wennberg D. Predictive risk project literature review. King's Fund, 2005. Available at www.kingsfund.org.uk/sites/default/files/field/field_document/predictive-risk-literature-review-june2005.pdf [Accessed 16 December 2021].
7. McCall N, Cromwell J, Urato C. Evaluation of Medicare care management for high cost beneficiaries (CMHCB) demonstration: Massachusetts General Hospital and Massachusetts General Physicians Organization (MGH). RTI Project Number 0207964.025.000.001. RTI International, 2010. Available at www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/mccall_mgh_cmhcb_final_2010.pdf [Accessed 16 December 2021].

8. Bottle A, Aylin P, Majeed A. Identifying patients at high risk of emergency hospital admissions: A logistic regression analysis. *J R Soc Med* 2006;99(8):406-14. doi: 10.1177/014107680609900818.
9. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366(6464):447-53. doi: 10.1126/science.aax2342.
10. Butler D, Clifford-Motopi A, Mathew S, et al. Study protocol: primary healthcare transformation through patient-centred medical homes—improving access, relational care and outcomes in an urban Aboriginal and Torres Strait Islander population, a mixed methods prospective cohort study. *BMJ Open* 2022 Sep 1;12(9):e061037. doi: 10.1136/bmjopen-2022-061037.
11. Butler DC, Agostino J, Paige E, et al. Aboriginal and Torres Strait Islander health checks: Sociodemographic characteristics and cardiovascular risk factors. *Public Health Res Pract* 2022;32(1):31012103. doi: 10.17061/phrp31012103.
12. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B* 1996;58(1):267-88. doi: 10.1111/j.2517-6161.1996.tb02080.x.
13. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33(1):1-22. doi: 10.18637/jss.v033.i01.
14. Ho TK. Random decision forests. In: *Proceedings of 3rd international conference on document analysis and recognition*, Vol. 1; 14-16 August 1995; Montreal, QC, Canada. IEEE, 1995: p. 278-82. doi: 10.1109/ICDAR.1995.598994.
15. Ke G, Meng Q, Finley T, et al. LighGBM: A highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst*. 2017;30:3146-54.

correspondence ajgp@racgp.org.au