POLAR Diversion: Using General Practice Data to Calculate Risk of Emergency Department Presentation at the Time of Consultation

Christopher Pearce[1] Adam McLeod[1] Natalie Rinehart[1] Jon Patrick[2] Anna Fragkoudi[1] Jason Ferrigi[1] Elizabeth Deveny[3] Robin Whyte[4] Marianne Shearer[5]

[1] Outcome Health, East Burwood, Victoria, Australia
[2] Health Language Analytics, Eveleigh, New South Wales, Australia
[3] South East Melbourne Primary Health Network, Melbourne, Australia
[4] Eastern Melbourne Primary Health Network, Box Hill, Victoria, Australia
[5] Gippsland Primary Health Network, Moe, Victoria, Australia

Address for correspondence Christopher Pearce, PhD, MFM, MBBS, GradDipPainMan, FRACGP, FACRRM, FAICD, FACHI, Outcome Health, 5 Lakeside Drive, East Burwood, Victoria 3344, Australia (e-mail: cpearce@outcomehealth.org.au).

Appl Clin Inform 2019;10:151–157.

Abstract

Objective This project examined and produced a general practice (GP) based decision support tool (DST), namely POLAR Diversion, to predict a patient's risk of emergency department (ED) presentation. The tool was built using both GP/family practice and ED data, but is designed to operate on GP data alone.

Methods GP data from 50 practices during a defined time frame were linked with three local EDs. Linked data and data mapping were used to develop a machine learning DST to determine a range of variables that, in combination, led to predictive patient ED presentation risk scores. Thirteen percent of the GP data was kept as a control group and used to validate the tool.

Results The algorithm performed best in predicting the risk of attending ED within the 30-day time category, and also in the no ED attendance tests, suggesting few false positives. At 0 to 30 days the positive predictive value (PPV) was 74%, with a sensitivity/recall of 68%. Non-ED attendance had a PPV of 82% and sensitivity/recall of 96%.

Conclusion Findings indicate that the POLAR Diversion algorithm performed better than previously developed tools, particularly in the 0 to 30 day time category. Its utility increases because of it being based on the data within the GP system alone, with the ability to create real-time "in consultation" warnings. The tool will be deployed across GPs in Australia, allowing us to assess the clinical utility, and data quality needs in further iterations.

Keywords
- electronic health record data
- primary care
- artificial intelligence
- machine learning
- risk prediction

Background and Significance

Approximately 70% of the total burden of disease in Australia is attributable to disease groups that could be either prevented or managed outside the inpatient system.[1] Reduction of avoidable hospital admissions is key to improving quality of life of patients and effectively managing expensive hospital resources. With activities at local, regional, and jurisdictional levels to reduce admissions,[2–4] a clinically proven mechanism to highlight patients at risk is essential.

Considerable research has gone into prediction of emergency presentation of patients,[5–9] predominantly using linear regression models. Numerous predictive algorithms/models have been developed over the years internationally with the aim to identify patients at high risk of emergency department (ED) presentation, admission, and readmission.[9–20]

received
August 4, 2018
accepted after revision
January 4, 2019

DOI https://doi.org/10.1055/s-0039-1678608.
ISSN 1869-0327.

These tools have had varying success, but have often struggled to show good sensitivity/recall over broader groups. In addition, they generally do not offer effective "real-time" information to inform general practitioners at the point in time care with patients. The POLAR Diversion project developed an automated algorithm based on family practice/general practice (GP) clinical and billing data to identify a score of ED presentation risk. The development of such a tool that can calculate patient alerts based on GP data alone can support practices to identify high-risk patients and where possible intervene to reduce their risk.

The Australian model of primary care puts GP at the center of care—all primary care physicians are GPs, and they are the gatekeeper to secondary care, including hospital care. Eighty-five per cent of the population see a GP each year.[21,22] The population has access to a government insurance scheme (Medicare) that subsidizes primary care visits and the public hospital system that is free at the point of care.[23] This means that any prevention or risk reduction tool must be aimed at GP to maximize impact. To support GPs in this, the government funds primary health networks (PHNs). As part of their programs the networks assist GPs in data quality and improvement activities.[23] This gave the project access to the needed pooled, deidentified data to use.

To achieve our goals, we chose to use a machine learning method. Machine learning has many variations, but they all share an important difference from the traditional statistical methods such as logistic regression or analysis of variance—the ability to make predictions on unseen data. To optimize the prediction accuracy, the methods generally do not attempt to produce interpretable models, but are rather designed to handle the large number of variables common in most big datasets, which under normal circumstances optimizes their prediction accuracy.[24] This technique was well suited to our large dataset.

Methods

Ethics

POLAR Diversion project was granted ethics approval from Monash University and Eastern Health Network for the collection, storage, and linkage of the datasets.

We targeted all conditions at risk for presentation, not just specific disease subgroups. Injury-based ED presentations were excluded from the dataset as these are not preventable at the point of consultation.

The detailed method has been published elsewhere.[25] The project essentially involved four phases.

- Five years of ED presentations were linked with GP data, to understand the GP journey of patients prior to an ED presentation.
- Mapping, grouping, and ranking of GP data to allow a consistent dataset to be fed into the algorithms.
- Using those data to develop a machine learning application to predict risk of an ED presentation, and test reliability with a set of unused data.
- Develop a decision support tool (DST) to be deployed in GP and receive feedback on tool performance.

GP data were extracted remotely from 50 practices from East Melbourne PHNs and linked with data from three local EDs from the same health service (Eastern Health) in the eastern suburbs of Melbourne. The ED data included date of attendance, length of stay, diagnosis, and demographic data to enable linking.

At the time of the GP data collection there were approximately 16 million deidentified GP patient records across 744,477 unique patients over the 5 years (2010–2015). A total of 12,448 of these patients had a linked ED presentation within the last year. As shown below, the sample used for the final algorithm decreased during the project as essential versus optional data were identified, removal of injuries within the ED data, etc. Once the model's parameters had been finalized the final data count for the algorithm included 17,067 GP visits, across 8,479 unique patients. The algorithm was also cross-checked across 29,892 GP visits for 29,185 unique patients for the non-ED sample. The sample is outlined in ►Table 1. An individual patient can have multiple visits. ►Table 2 gives the timing of the individual visit according to the time between the last GP visit and the ED presentation. It counts multiple visits against the index GP visit—hence the total number is larger than the ED visit count.

Significant data preparation work was undertaken: mapping diagnoses to a standard terminology (SNOMED-CT-AU); medications to the World Health Organization's Anatomic

Table 1 Algorithm sample from general practice

Criteria	GP visits	Unique patient count
All data supplied	16,305,096	744,477
Those with a linked ED visit	37,789	21,376
Those with a linked ED visit within 1 year	26,691	12,448
Those with a linked ED visit within 1 year and no injury	20,213	10,610
Those with a linked ED visit within 1 year, no injury, with adequate data fields	17,067	8,479
Those without a linked ED visit used for the non-ED sample	150,000	144,490
Those without a linked ED visit used for the non-ED sample with adequate data fields	29,892	29,185

Abbreviations: ED, emergency department; GP, general practice.

Table 2 Intervals between last GP visit and ED presentation

Time (days)	Total visits
0–30	15,051
31–90	5,320
91–180	3,006
180–365	2,989

Abbreviations: ED, emergency department; GP, general practice.

and Therapeutic Classification. A clinician (C.P.) then ranked all drug doses to usual dose, high or low dose, according to standard dosing recommendations. Pathology was mapped to a clinical model relevant for GP (i.e., grouping all hepatitis testing), as well as a series of severity mappings for key measures such as blood pressure (BP), body mass index (BMI), etc. The model was then built with these given attributes. The tool used 21 *group attributes* (i.e., medication), including 52 *relational attributes* (i.e., medication dose and/or frequency) to inform the algorithm. This means many thousands of variables were being taken into account to build the risk score. Further to this process, given the intended use for a DST within GP, the data were divided into the historical data (prior to the index consultation) and the data at the last (index) consultation. In addition, time relevance needed to be set for each of the variables, for example, medications older than 2 years in the GP record, was not considered. This allowed the data inputs to be ranked, highlighting specific high-risk groups of diagnoses, medications, or pathology as key factors that may further contribute to the machine learning. For each patient with multiple hospital attendances, each attendance was treated separately.

The project used a machine learning program developed specifically for this purpose by Health Language Analytics (J.P.). The major algorithm building tasks included the following.

- Building a coherent representation of the patient records suited for computing a predictive model.
- Testing a variety of combinations of attributes for the best results.
- Converting the many attributes available into domain ranges that were relevant to the task.
- Testing many class configurations around 30-day, 90-day, 180-day, 365-day, and post-1-year attendances.
- Devising representations of the various time lapses between the GP visits of patients;

We kept back 13% of the data for use as a testing set, according to standard practice. These data were then parsed through the machine learning program to determine the tool's ability to accurately identify both patients who had attended an ED and those who had not.

Results

Model Development

After extensive data review, the data were split between the data entered at the current visit and a historical view. This was to deal with the relevance of certain types of information for a current ED presentation. The notion was to separate the more recent patient information to the GP from previous visits, which were then collapsed into the one "*historical visit*." The criteria for key attributes were specifically designed and are presented in ►**Table 3** below.

In addition to this, extra mapping was laid over the data including risk groups for key measurements (i.e., BMI, BP, blood sugar levels, cholesterol, falls, and temperature) and a medication risk grouping where particular medications were given a higher risk score than others (i.e., medications for cardiac issues or chemotherapy were scored at a higher risk than acne medication). These extra mappings allowed for another dimension of understanding of the data to be

Table 3 Final variable listing in current and historical visits

Attributes	Time categories
Clinical fields	
Current diagnoses	Current visit and ACTIVE diagnoses
Historical diagnoses	Up to 10 years—not including current diagnosis information
Current immunization	Current visit
Historical immunization	Within past 5 years—not including current visit
Current prescriptions	Within last 8 months
Historical prescriptions	Between 9 and 24 months
Current pathology test	Current visit
Current pathology result	Current visit
Historical pathology test	Within last 12 months—not including last visit
Historical pathology result	Within last 12 months—not including last visit
Current measurement	Current visit
Historical measurement	Within past 5 years—not including current visit
Billing details	Any
Demographic and other patient information	
Alcohol usage	Last recorded
BP recorded	Last recorded and rated
Care goal	Last recorded
Reaction (allergies)	Last visit
Historical reaction	All information apart from last visit
Tobacco	Last recorded
Age	Last visit
Sex	Last visit
Department of Veterans Affairs status	Last visit
Pension status	Last visit
Aboriginal status	Last visit

Abbreviation: BP, blood pressure.

worked into the model. The data were subjected to 10-fold cross-validation on a support vector machine, identifying the precision and recall for each class.[25]

GP Data Item Importance

The importance of items was reviewed across the final model. ►Fig. 1 shows the relative importance of the many variables in the model. ►Appendix A explains the individual feature importance categories. Some of the key variables are diagnoses and pathology results, with historical diagnosis being the most significant followed by historical pathology result, with less, but still apparent input from medications variables, both current and historical sitting on the second tier of significance. Individual patient demographics such as age, gender, and pension status were less important than the clinical information of patients.

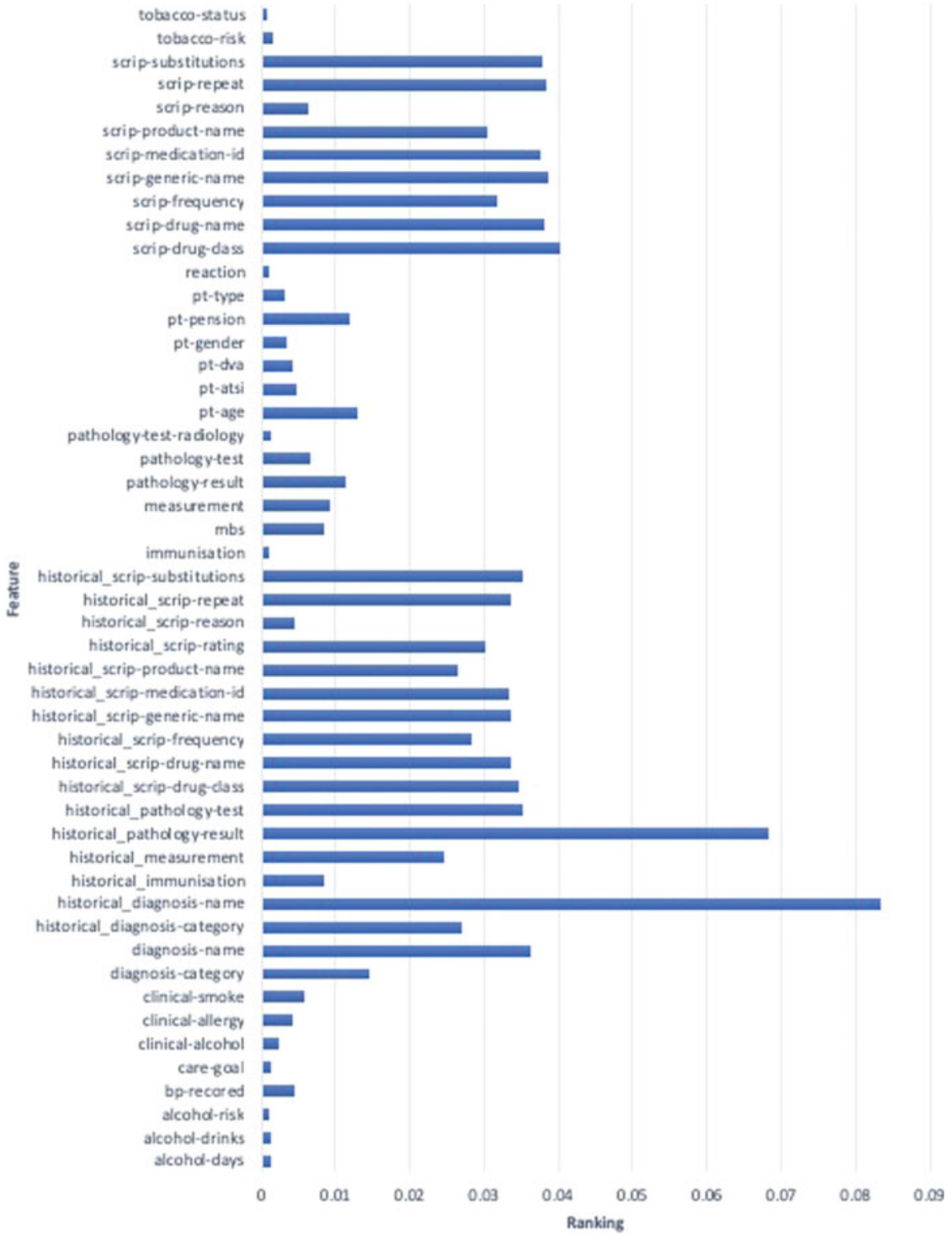

Fig. 1 Feature importance for the final model.

Table 4 Algorithm outcomes

Algorithm	Positive predictive value/precision (%)	Sensitivity/ recall (%)
0–30 days	74	68
31–365 days	37	10
No ED attendance	82	96

Abbreviation: ED, emergency department.

Eighty-six per cent of visits had some value for previous diagnoses, with the most common including chronic obstructive pulmonary disease, bone joint disease, diabetes, cancer, coronary heart disease, asthma, gastroenteritis, stroke, influenza, hypertension, anxiety, depression, and hepatitis.

The most frequent diagnoses/reasons for visits at the time of the visit were the following.

- Hypertension.
- Upper respiratory tract infection.
- Asthma.
- Depression.
- Bronchitis.
- Tonsillitis.
- Urinary tract infection.
- Otitis media.
- Gastroenteritis.
- Review.

Risk Prediction

When parsed through the testing set, the results were very encouraging. Overall, the algorithm performed best in the 0 to 30 day time category with a 73.7% precision score. The 31 to 365 day time category had a positive predictive value (PPV) of 36.8% and the "no ED attendance" category showing a precision score of 82.3%.

As shown below in ►Table 4, the algorithm performed best in the 0 to 30 day time category across both the PPV and sensitivity/recall scores. This was also evident in the non-ED attendance tests indicating few false positives.

Conclusion

The focus of this project was to develop a DST on a tool that uses GP data generated at the time of consultation to increase its utility, in terms of a program that can deliver real-time warnings. It is designed to be automated to run over an entire patient population. However, there are several outstanding questions. Not the least of which is: what is the likely benefit of such warnings to clinicians and patients? The DST can generate a list of the characteristics used to make its decision. It is then up to the individual patient and doctor to decide how best this can be approached. In that sense, this program is designed to be an aid to the patient–doctor relationship, not a replacement of the consultation. Post-deployment analysis and monitoring will be crucial to understanding the success (or otherwise) of artificial intelligence in the primary care context. The tool also gives feedback on data missing from the record—with which it could have made a better decision. By prompting the GP to complete the missing information in the patient's record the tool acts as a data quality improvement tool as well.

Overall, the POLAR Diversion algorithm performed better than most existing tools based on GP data, particularly in the 0 to 30 day time category, and in the sensitivity/recall scores. So POLAR at 0 to 30 days (74/68) is comparable with Q admissions (73/70 over 1 year),[15] and exceeds PARR + +(84/1)[26] and Peony (67/4)[13]—the most widely reported studies prior to this one. This may be due to a combination of reasons such as access to the data, significant mapping and cleaning of data, and the first use of machine learning in Australia for this context (in place of the commonly used regression analyses in other models). Machine learning allows for the inclusion of a large breadth of data and types of data into modeling and is becoming more common with health and biomedical research globally. Whereas many of the previous models included key "*risk factors*" for consideration such as chronic conditions, key specific medications or pathology outcomes, and recent hospital attendances, the POLAR Diversion algorithm included a large range of GP data, thereby, increasing its possible reach.

The removal of the "*injury*"-based ED attendances from the modeling ensured that it was trained on more predictable ED presentations with real and trackable health concerns. In most cases, injuries cannot be predicted unless they are due to a medical condition (i.e., falls in the aged). But even so—the DST should predict the risk of a fall in the next 30 days—once the fall has occurred, no preventive action can be taken. The removal of these events from the model will strengthen its connectivity to relevant GP patient data. The tool was generated on linked ED/GP data from the eastern suburbs of Melbourne, and therefore is most applicable there. However, the nature of machine learning is that the tool can evolve over time, as more data (from a wider variety of GP settings) are added to the model. Local factors (such as rurality) can be factored into the model as well, as time and resources permit. Nevertheless, any such a tool is subject to the data it is fed with, and this particular model is static—in that it was trained on a limited dataset. By definition, the intent of the program is to reduce the admission rate, thereby changing the underlying parameters. Our decision is to deploy the model into participating GPs (500–1000 across the country) as an "*advanced decision support*" tool, providing advice to the clinician at the point of consultation. This will have to be monitored and assessed to determine what use clinicians make of the tool, and how to improve its implementation. At the same time, we have received funding to greatly expand the data pool—retraining the model on data from hospitals and GPs across other parts of Australia, rather than one geographical area in one city.

Machine learning and its application to medicine is subject to much hype and fear, uncertainty, and doubt.[27,28] Overall, this study has demonstrated the value of using linked data and modern computing tools to generate a machine learning model that has a high rate of predicting the risk of admission based on GP electronic health record data alone. This allows it to be deployed at the point of care,

to maximize its effectiveness in influencing care provision. It represents a possible future where such techniques can aid both patients and doctors to use the vast amounts of data available to improve their care options.

Multiple Choice Questions

1. In considering a machine learning tool for primary care it is important to:
   a. Deal with information from a single clinical system.
   b. Have data divided into variable with attributes.
   c. Use linked data with hospital care.
   d. Place geographic limitations on the data.

   Correct Answer: The correct answer is option b.

2. In considering the results, the precision score represents:
   a. The degree with which the tool measures the risk of admission.
   b. The rate of agreement in a cohort of data withheld to test the accuracy.
   c. The relative quality of the data.
   d. The success of the tool in a real-world environment.

   Correct Answer: The correct answer is option a.

Protection of Human and Animal Subjects

No human subjects were involved in this project.

Funding

This research was undertaken with the generous assistance from the HCF Research Foundation. The foundation exerted no influence over the design or the progress of the study.

Conflict of Interest

None declared.

References

1 Begg S, Vos T, Barker B, Stevenson C, Stanley L, Lopez AD. The burden of disease and injury in Australia 2003. PHE 82. Canberra: AIHW; 2007
2 Harrison MJ, Dusheiko M, Sutton M, Gravelle H, Doran T, Roland M. Effect of a national primary care pay for performance scheme on emergency hospital admissions for ambulatory care sensitive conditions: controlled longitudinal study. BMJ 2014;349:g6423
3 Billings J, Georghiou T, Blunt I, Bardsley M. Choosing a model to predict hospital admission: an observational study of new variants of predictive models for case finding. BMJ Open 2013;3(08):e003352
4 Adelaide PHN. 2018. Available at: https://www.adelaidephn.com.au/assets/Core_Operational_and_Flexible_AWP_2018_FINALDoH_version270918_Publicver.pdf. Accessed January 21, 2019
5 Reis BY, Mandl KD. Time series modeling for syndromic surveillance. BMC Med Inform Decis Mak 2003;3(01):2
6 Jones SS, Thomas A, Evans RS, Welch SJ, Haug PJ, Snow GL. Forecasting daily patient volumes in the emergency department. Acad Emerg Med 2008;15(02):159–170
7 Schweigler LM, Desmond JS, McCarthy ML, Bukowski KJ, Ionides EL, Younger JG. Forecasting models of emergency department crowding. Acad Emerg Med 2009;16(04):301–308
8 Wargon M, Casalino E, Guidet B. From model to forecasting: a multicenter study in emergency departments. Acad Emerg Med 2010;17(09):970–978
9 Crilly JL, Boyle J, Jessup M, et al. The implementation and evaluation of the patient admission prediction tool: assessing its impact on decision-making strategies and patient flow outcomes in 2 Australian hospitals. Qual Manag Health Care 2015;24(04):169–176
10 Reuben DB, Keeler E, Seeman TE, Sewall A, Hirsch SH, Guralnik JM. Development of a method to identify seniors at high risk for high hospital utilization. Med Care 2002;40(09):782–793
11 Asplin BR, Flottemesch TJ, Gordon BD. Developing models for patient flow and daily surge capacity research. Acad Emerg Med 2006;13(11):1109–1113
12 Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. BMJ 2006;333(7563):327–330
13 Donnan PT, Dorward DW, Mutch B, Morris AD. Development and validation of a model for predicting emergency admissions over the next year (PEONY): a UK historical cohort study. Arch Intern Med 2008;168(13):1416–1422
14 Essex Strategic Health Authority Combined Predictive Model. Final Report. 2006. UK. Available at: https://www.kingsfund.org.uk/sites/default/files/field/field_document/PARR-combined-predictive-model-final-report-dec06.pdf. Accessed January 21, 2019
15 Hippisley-Cox J, Coupland C. Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. BMJ Open 2013;3(08):e003482
16 HARP W. Summary. HARP risk calculator. 2009. Available at: https://www.adma.org.au/clearinghouse/doc_details/11-western-harp-risk-calculator.html. Accessed January 21, 2019
17 Howell S, Coory M, Martin J, Duckett S. Using routine inpatient data to identify patients at risk of hospital readmission. BMC Health Serv Res 2009;9(01):96
18 Khanna S, Boyle J, Good N. Precise prediction for managing chronic disease readmissions. Conf Proc IEEE Eng Med Biol Soc 2014;2014:2734–2737
19 Rana S, Tran T, Luo W, Phung D, Kennedy RL, Venkatesh S. Predicting unplanned readmission after myocardial infarction from routinely collected administrative hospital data. Aust Health Rev 2014;38(04):377–382
20 Gildersleeve R, Cooper P. Development of an automated, real time surveillance tool for predicting readmissions at a community hospital. Appl Clin Inform 2013;4(02):153–169
21 Beilby J, Furler J. General practitioner services in Australia. In: Pegram R, ed. General Practice in Australia: 2004. Canberra: Australian Government Publishing Service, Australian Department Of Health and Ageing; 2005:128–213
22 Duckett S, Willcox S. The Australian Health Care System. 5th ed. Oxford: Oxford University Press; 2015
23 Pearce C, Shearer M, Gardner K, Kelly J, Xu TB. GP Networks as enablers of quality of care: implementing a practice engagement framework in a general practice network. Aust J Prim Health 2012;18(02):101–104
24 Beam AL, Kohane ISJJ. Big data and machine learning in health care. JAMA 2018;319(13):1317–1318
25 Pearce CM, McLeod A, Patrick J, et al. Using patient flow information to determine risk of hospital presentation: protocol for a proof-of-concept study. JMIR Res Protoc 2016;5(04):e241
26 Johnson S. Risk Stratification and Next Steps with DH Risk Prediction tools. UK Government. London: Department of Health; 2011
27 Pearce C, Haikerwal MC. E-health in Australia: time to plunge into the 21st century. Med J Aust 2010;193(07):397–398
28 Yan J, Hawes L, Turner L, Mazza D, Pearce C, Buttery J. Antimicrobial prescribing for children in primary care. J Paediatr Child Health 2018;55(01):54–58

Appendix A Explanation of feature importance categories

Tobacco status	Current or past tobacco use
Tobacco risk	Risk rating based on tobacco use
Scrip substitution	If brand name has been substituted for different generic brand
Scrip repeat	Number of repeats on prescription
Scrip reason	Reason for prescription if recorded
Scrip product name	Brand name of drug prescribed
Scrip medication ID	Specific drug identification identifier
Scrip generic name	Generic drug brand
Scrip frequency	How often drug is to be taken/used
Scrip drug name	Chemical name
Scrip drug class	Therapeutic class
Reaction	Identified allergic reaction
Patient type	
Patient pension	Presence or absence of pension and type (age/disability)
Patient gender	
Patient DVA	Department of Veteran Affairs eligible
Patient ATSI	Identifies as Aboriginal or Torres Strait Islander descent
Patient age	In years
Pathology test: radiology	Presence of a radiology test and type
Pathology test	Presence of a pathology test and type
Pathology result	Results (with attributes of high or low, where given)
Measurement	Class of measurements such as weight, height, peak flow etc..
MBS	Medicare Benefit Schedule—billing details.
Immunization	Presence of immunizations
Historical scrip substitution	As above, but in the historical visits section
Historical scrip repeat	As above, but in the historical visits section
Historical scrip reason	As above, but in the historical visits section
Historical scrip product name	As above, but in the historical visits section
Historical scrip medication ID	As above, but in the historical visits section
Historical scrip generic name	As above, but in the historical visits section
Historical scrip frequency	As above, but in the historical visits section
Historical scrip drug name	As above, but in the historical visits section
Historical scrip drug class	As above, but in the historical visits section
Historical pathology test	As above, but in the historical visits section
Historical pathology result	As above, but in the historical visits section
Diagnosis name	Recorded diagnosis
Diagnosis category	SNOMED code and grouping of diagnosis
Clinical smoke	Smoking advice recorded
Clinical allergy	Recorded allergies
Clinical alcohol	Recorded alcohol use
Care goal	Presence and type of care goals in care plan
BP recorded	BP recorded in system at the time of visit
Alcohol risk	Calculated from alcohol recording
Alcohol drinks	Standard drinks per day
Alcohol days	Number of days per week.

Abbreviations: BP, blood pressure; ID, identity document.