Novel machine learning model for predicting multiple unplanned hospitalisations

Paul Conilione [1], Rebecca Jessup [2,3] Anthony Gust [1]

¹Digital Health, Northern Hospital, Epping, Victoria, Australia
²Staying Well, Northern Hospital, Epping, Victoria, Australia
³School of Allied Health, Human Service and Sport, La Trobe University, Melbourne, Victoria, Australia

Correspondence to
Dr Paul Conilione;
paul.conilione@nh.org.au

ABSTRACT

Background In the Australian public healthcare system, hospitals are funded based on the number of inpatient discharges and types of conditions treated (casemix). Demand for services is increasing faster than public funding and there is a need to identify and support patients that have high service usage. In 2016, the Victorian Department of Health and Human Services developed an algorithm to predict multiple unplanned admissions as part of a programme, Health Links Chronic Care (HLCC), that provided capitation funding instead of activity based funding to support patients with high admissions.

Objectives The aim of this study was to determine whether an algorithm with higher performance than previously used algorithms could be developed to identify patients at high risk of three or more unplanned hospital admissions 12 months from discharge.

Methods The HLCC and Hospital Unplanned Readmission Tool (HURT) models were evaluated using 34 801 unplanned inpatient episodes (27 216 patients) from 2017 to 2018 with an 8.3% prevalence of 3 or more unplanned admissions in the following year of discharge.

Results HURT had a higher AUROC (84%, 95% CI 83.4% to 84.9% vs 71%, 95% CI 69.4% to 71.8%) than HLCC, that was statistically significant using Delong test at p<0.05.

Discussion We found features that appear to be strong predictors of admission risk that have not been previously used in models, including socioeconomic status and social support.

Conclusion The high AUROC, moderate sensitivity and high specificity for the HURT algorithm suggests it is a very good predictor of future multi-admission risk and that it can be used to provide targeted support for at-risk individual.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Case-finding algorithms for identifying patients at risk of unplanned readmissions traditionally have focused on detecting one or more admissions over a 30-day to 365-day period from discharge. This study focuses on patients who have more frequent admissions, and aims to predict three or more unplanned admissions within 365 days of an index admission. This allows for better targeting of patients that would otherwise use more resources. Accurately predicting those who represent the highest hospital use is likely to lead to greater healthcare cost savings.

WHAT THIS STUDY ADDS

⇒ This study presents an algorithm for identifying patients at risk of three or more unplanned admissions using not only clinical information, socioeconomic indicators and living arrangements, but also a novel cascading chronic condition feature.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study demonstrates the importance of using a mixture of clinical, demographic and activity-based features for predicting patient outcomes. Our algorithm outperformed a similar algorithm that used existing weighted scoring approaches. The study demonstrates that machine learning-based methods for identifying patients who would benefit from targeted intervention have great potential in improving health system sustainability.

BACKGROUND

Potentially preventable admissions include any hospitalisations for acute and chronic conditions that may have been avoided with earlier intervention and rehospitalisation within 30 days of discharge due to inadequate discharge and/or follow-up.[1] In many high-income countries, potentially preventable hospitalisations have become an indicator of health system performance.[2] Reported rates of preventable hospitalisation range from 5% to 79%.[3] This wide range reflects not only differences in the definition of preventable admissions, but also geographical and socioeconomic differences in population composition.[4] For hospitals, potentially preventable admissions increase hospital demand, lead to bed blocking and patient flow issues, and in Australia account for 10% of all occupied beds and more than 748000 admissions per year.[5]

The cost of providing healthcare in most high-income countries is considered to be unsustainable and will likely be unaffordable by 2050 in the absence of major reforms.[6] Identifying and averting these preventable hospitalisations is important for not only improving individual health outcomes but

in controlling burgeoning healthcare expenditure. Case finding algorithms to identify those at highest risk of preventable hospitalisations are emerging as a key initiative that may allow for targeted care to prevent deterioration and future admissions.[7] A highly sensitive and specific case-finding algorithm should be able to identify only those patients most likely to have high future healthcare costs or hospital resource use.

Internationally, there is a growing body of literature on algorithms that aim to predict the likelihood of future admissions using different models, including the traditional logistic regression model and survival analysis and more recently popular modelling using machine learning techniques.[8] Many approaches focus on patients at risk in specific disease categories such as chronic obstructive pulmonary disease (COPD),[9] stroke/Transient Ischaemic Attack (TIA),[10] diabetes[11] or heart failure.[12] Others focus on unplanned readmissions, usually within 30 days of discharge, for any-cause using non-linear models,[13] gradient boosted decision trees[14] or artificial neural networks.[15]

In 2016, the Victorian Department of Health and Human Services (DHHS) initiated the HealthLinks Chronic Care programme (HLCC) that provided an alternative capitated funding model for patients with chronic and complex health conditions who were at high risk of multiple unplanned admissions. A key component of the programme was the use of a predictive algorithm called the HLCC model. The HLCC model uses an index unplanned admission as a triggering event and then combines diagnostic information from that admission and demographic information to create a 'risk score' for the probability of another three or more admissions in the next 12 months. Patients who score above a threshold value determined by logistic analysis of historical data are eligible to be included in the HLCC programme, receiving targeted preventative care.[16] The HLCC risk score was found to have a sensitivity of 41% and specificity of 78% over the 2-year evaluation across five participating Victorian health service providers.[17] The low sensitivity suggests that there are potentially many patients who would benefit from targeted intervention who are not being identified by this algorithm and the moderate specificity suggests that efforts with targeted intervention was wasted on some individuals who would not have gone on to have a preventable admission. This paper describes the development, and content, of a machine learning case-finding prediction tool with a higher sensitivity and specificity for identifying patients that are at high risk of all-cause potentially avoidable admissions within 12 months of discharge in an Australian setting.

METHODS
Setting
This was a single-centred study based at Northern Health (NH). NH is the major provider of acute (410 beds), subacute (251 beds) and ambulatory specialist services in Melbourne's north. Residents originate from more

than 184 countries, speak more than 106 languages and have lower levels of income, educational attainment and health literacy, and higher rates of unemployment than state averages.[18] The emergency department at NH is the busiest in the state with over 100,000 presentations per year.[19]

Study design and data sources
Participants
Eleven years of historical NH acute Inpatient (IP) emergency admitted episodic level data was used from 1 July 2008 to 30 June 2019 to build and test a new model that we named the HURT (Hospital Unplanned Readmission Tool) model. Outpatient (OP) and emergency department (ED) data were also linked to the unplanned IP activity. In addition, the Index of Relative Socio-economic Advantage and Disadvantage (IRSEAD) from the Australian Bureau of Statistics (ABS) Socio-Economic Indexes for Areas (SEIFA) data set were linked to patient's residential postal address.

An unplanned admission is an unexpected or sudden health issue or event that results in an emergency admission. We only included acute IP episodes where the patient was 18 years or older at admission, the admission was not related to mental health, obstetrics, oncology or renal dialysis and the patient did not die during the episode. If there were any records that contains missing values then they were discarded.

Table 1 presents the summary statistics of the data used including demographic information and the features used for the final model. Where the percentages are the proportion of separations with the given flag. The features are defined in table 2 later in the paper. These data were included as DHHS and other jurisdictions have these data readily available. Data such as pathology and pharmacy were not included as the DHHS does not have this.

Patient and public involvement
Patients and the public were not involved in the design of this work.

Variable selection for the HURT
Variable (feature) selection is a manual or automatic process by which variables that have the highest impact on model performance (in this case prediction of future unplanned admissions) are selected and variables that do not help learning are discarded.

The Boruta R package was used to develop the HURT. Boruta R uses a novel feature selection algorithm that finds all relevant variables, where relevant means variables that are found to be associated with unplanned emergency admissions. Boruta can use a range of decision trees to derive the importance of each feature. Extreme Gradient Boosting (XGBoost) was used to measure the feature importance in the Boruta algorithm with 200 maximum runs to ensure feature importance was fully resolved.

We also created a novel feature which we called a cascading chronic condition flag. If a patient was coded

Table 1 Summary statistics of the 11 years of data with mean value or percentage of separations with the relevant flag

Property	All	Readmitted <3	Readmitted ≥3
No of separations	206 714	192 679	14 035
No of patients	125 743	125 258	4730
Female %	50.3%	50.5%	47.8%
Average admission age	55.9	55	67.7
Admission age 61–90 flag	41.5%	39.7%	66.7%
Admission age 90+ flag	2.7%	2.6%	3.8%
Chronic condition (cascading) COPD flag	4.5%	3.6%	17.4%
Chronic condition (cascading) disorder due to tobacco flag	6.4%	5.2%	22.7%
Chronic condition (cascading) heart failure flag	5.4%	4.3%	19.5%
Complexity ≥3 flag	27.4%	25.7%	50.1%
Failed to attend ratio OP 365 days	8.7%	7.7%	21.5%
Marital status flag	22.7%	21.5%	39.2%
No OP bookings in past 365 days flag	65.5%	68.1%	30.3%
No of HIPs (any group)	0.22	0.17	0.83
No of non-admitted ED presentations	0.5	0.4	1.5
No of OP attended past year	1.5	1.3	4
Potential avoidable emergency admission flag	25.5%	24.3%	42.4%
IRSEAD decile within Australia	4.6	4.6	4
Total LOS unplanned episodes 180 days	4.6	4.1	10.7
Total LOS unplanned episodes 365 days	5.4	4.7	14.9
Total no unplanned eisodes past 365 days	1.6	1.5	3.6
Usual accommodation agecare, alone flag	2.6%	2.4%	6.2%

COPD, chronic obstructive pulmonary disease; ED, emergency department; HIP, Health Independence Program; IRSEAD, Relative Socio-economic Advantage and Disadvantage; LOS, Length of Stay; OP, emergency department.

with a chronic condition, all subsequent episodes were also flagged for this chronic condition (ie, if a patient was diagnosed with COPD, all following episodes would be flagged with this condition, when previously this would not occur if their admission had been for a different reason). Hence, this information can be used by the machine learning (ML) modelling to help predict future unplanned admissions.

Weighting variable importance

Over the past decade, 'black box' machine learning algorithms have been increasingly used in critical decision-making processes. However, because it is unclear or unknown how the machine learning algorithm decides there have been reports of adverse results in some fields.[20]

To overcome this problem, we used interpretable models that allow for an understanding of why the machine learning algorithm makes particular decisions on individual cases. The SHAP (Shapley Additive exPlanations) score was used as it assigns each variable an importance value for each decision outcome. The SHAP score can then be visualised to illustrate how the decision tree-based machine learning is making a given decision in an interpretable manner.[21]

Training and optimisation of the model

HURT is trained and tested on historical data where we know in advance if a patient had three or more unplanned admissions 1 year from the index admission. We define this as the 'target' for the model to be trained and tested on.

The XGBoost machine learning algorithm uses an ensemble (collection) of weak decision trees that are sequentially created to progressively improve (ie, boosting) the learning performance.[22] This has the advantage of quick training and has been shown to perform well on unseen test data. It also has the advantage of being able to handle unbalanced data where there are fewer positive cases (ie, patient returned three or more times in the future) compared with the negative cases (ie, patient did not return three or more times).

Like most machine learning algorithms, XGBoost has a set of training parameters that impact the final model performance. The parameter values that maximise performance cannot be determined by analysing the data only. These can only be found by trying different training parameters and measuring the model performance.[23] Hence, we performed hyperparameter optimisation by using a simple grid-search over a range of parameter

Table 2 List of variables in final model for predicting three or more unplanned admissions

Feature	Data type	Description
Admission age 61–90 flag	Binary	1 if Admitted age 61–90, 0 otherwise
Admission age 90+ flag	Binary	1 if admitted age ≥90, 0 otherwise
Chronic condition (cascading) COPD flag	Binary	Cascading chronic condition flag for COPD
Chronic condition (cascading) heart failure flag	Binary	Cascading chronic condition flag for heart failure
Chronic condition (cascading) disorder due to tobacco flag	Binary	Cascading chronic condition flag for disorder due to tobacco use
Complexity ≥3 flag	Binary	1 if number of body systems treated is 3 or more,[28] 0 otherwise.
Marital status flag	Binary	1 if divorced, widowed or separated, 0 otherwise
Total unplanned episodes past 365 days	Integer	In past 365 days prior to discharge
No of non-admitted ED presentations 365 days	Integer	Calculated for past 365 days from discharge
No of HIPs (any group)	Integer	Number of enrolments in any HIP group for past 180 days
No of OP attended	Integer	In past 365 days from IP discharge
Failed to attend ratio OP 365 days	Float	Calculated for past 180 and 365 days from discharge
No OP bookings flag past 365 days flag	Binary	1 if no OP bookings in past year, 0 otherwise
Usual accommodation age care, alone flag	Binary	1 if patient is living in age care facility or living alone, 0 otherwise
Potential avoidable emergency admission flag	Binary	1 if patient had ICD10 diagnosis code from,[29] 0 otherwise
IRSEAD Decile Within Australia	Integer	ABS Index of Relative Socio-economic Advantage and Disadvantage by postcode (0-low, 10-high)
Total LOS unplanned episodes 180 days	Integer	Calculated for past 180 days from discharge
Total LOS unplanned episodes 365 days	Integer	Calculated for past 365 days from discharge

ABS, Australian Bureau of Statistics; COPD, chronic obstructive pulmonary disease; ED, emergency department; HIP, Health Independence Program; ICD10, International Statistical Classification of Diseases and Related Health Problems, 10th Version; IP, inpatient; IRSEAD, Index of Relative Socio-economic Advantage and Disadvantage; LOS, Length of Stay; OP, outpatient.

values. The optimal XGBoost parameters where Eta=0.05, Max Depth=4, Gamma=0, Colsample_bytree=1, min_child_weight=2, Subsample=0.5, Nrounds=400.

The 11 years of historical data were divided into training and testing phases. The optimal model parameter values that produced the highest area under the reciever operator curve (AUROC) performance using 10-fold cross-validation on the training data (9 years, 171 913 separations, 98 527 patients) were used. Testing was performed on the episodes that were discharged in 2017–18 (1 year, 34 801 separations, 27 216 patients), but 2018–2019 data were needed to count the unplanned admissions up to 1 year from 2017 to 18 discharge. The training and testing phase were performed by using the Caret R-package.[24]

Final variables selected for the HURT

Table 2 provides an overview of the final 18 features selected for HURT from an initial set of 199 features. The definition of all features tested are available as online supplemental material 1.

The performance of the HURT was assessed retrospectively by calculating the AUROC, sensitivity which is the percentage of separations where the patient was correctly predicted to have three or more potentially avoidable admissions in the 12 months following their discharge. We

also assessed the specificity of the model (the percentage of patient separations incorrectly predicted to have three or more unplanned admissions). The higher the sensitivity of the model, the more patients correctly identified and the less that will be 'missed' and have a potentially preventable readmission.

Comparison

The primary comparison of our algorithm is with the DHHS HLCC algorithm using AUROC, sensitivity and specifity. We also compare the decisions made by HURT and HLCC on the same separations and illustrate the differences by a Venn diagram. Of particular interest is the false-positives (FP) where a patient is falsely flagged as returning three or more times and will be offered support services. This means resources may be used for patients that were not going to return. The false-negative (FN) cases are of concern as these patients are not flagged, and will not be offered extra services, thus returning three or more times since discharge. This places a strain on hospital resources that could have been reduced but more importantly potential missing patients that may have deteriorated.

Given the lack of existing research using three or more unplanned admissions within 1 year of discharge. We also applied the previously described methodology
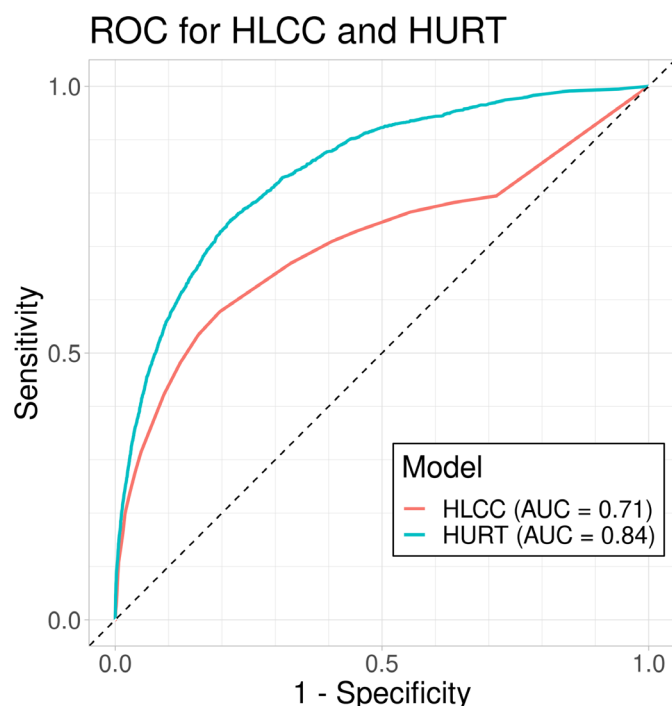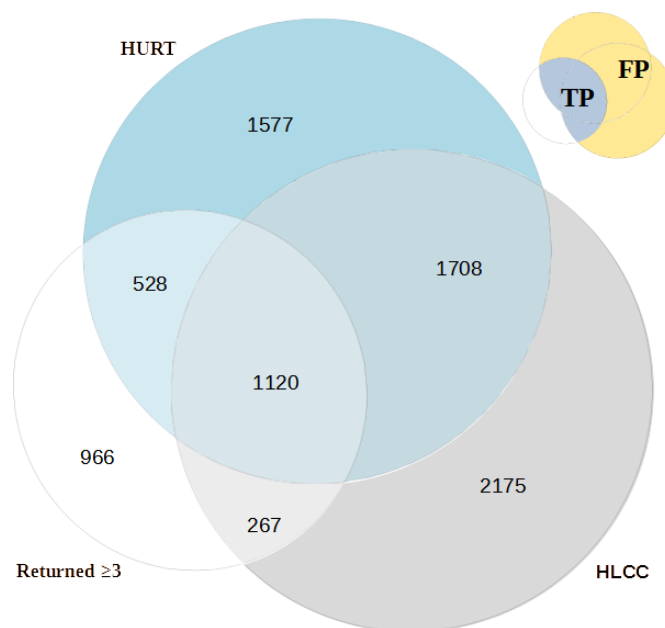
Figure 1 ROC test performance of HLCC and HURT models predicting three or more unplanned admissions. HLCC, HealthLinks Chronic Care; HURT, Hospital Unplanned Readmission Tool; ROC, receiver operating characteristic.

Figure 2 Venn diagram of number of separations that were predicted to return by machine learning, HLCC and overlap with the number of separations that actually returned three or more times. FP, false-positive; HLCC, HealthLinks Chronic Care; HURT, Hospital Unplanned Readmission Tool; TP, true-positive.

to predicting if a patient has one or more unplanned admissions within 1 year. Our model is compared with other case-finding algorithms that predict one or more unplanned admission over a year using weighted scores[25][26] and machine learning.[27] Even though this was not the focus of the research, it provided some reassurance of the methodology.

RESULTS

The optimised HURT model had a final test AUROC of 84% (95% CI 83.4% to 84.9%), while HLCC had an AUROC of 71% (95% CI 69.4% to 71.8%) (figure 1). The difference between HURT and HLCC ROC was statistically significant with Z=−22.6, p<0.001 (Delong test).

Using the confusion matrix in table 3, the HURT algorithm had a sensitivity of 57%, while HLCC had a sensitivity of 48%. The 9% difference was statistically significant with χ^2=85.03, p<0.001 (McNemar test). The

HURT algorithm achieved 90% specificity, while HLCC had a specificity of 88%. The 2% difference was statistically significant with χ^2=94.99, p<0.001 (McNemar test).

The Venn diagram in figure 2 provides an overview of the number of hospital unplanned admissions that were predicted by each of the HLCC and HURT models in terms of true-positive (TP) and FP cases. The number of separations that were predicted correctly (ie, the overlap between 'returned≥3', HURT and HLCC) are TP cases (HURT: 528, both: 1120, HLCC: 267). Where HURT has 261 more separations correctly classified compared with HLCC. While the FP cases (HURT: 1577, both: 1708, HLCC: 2175) show the HURT has 598 fewer FPs compared with HLCC. Both models missed 966 positive cases.

The 18 most important variables for predicting admission can be grouped into three: demographics (particularly age and marital status), medical conditions (complexity and cascading chronic conditions, in particular COPD and chronic cardiac failure) and

Table 3 Comparison of the HURT and HLCC algorithms in identifying patients at risk of three or more unplanned readmissions

Algorithm	HLCC		HURT	
Prediction	Readmitted ≥3	Readmitted <3	Readmitted ≥3	Readmitted <3
Readmitted ≥3	1387 (TP)	1494 (FP)	1648 (TP)	1233 (FP)
Readmitted <3	3883 (FN)	28 037 (TN)	3285 (FN)	28 635 (TN)
Result	Sensitivity 48%	Specificity 88%	Sensitivity 57%	Specificity 90%

FN, false-negative; FP, false-positive; HLCC, HealthLinks Chronic Care; HURT, Hospital Unplanned Readmission Tool; TN, true-negative; TP, true-positive.

Figure 3 SHAP plot of impact of each feature on decision of XGBoost model. COPD, chronic obstructive pulmonary disease; ED, Extreme Gradient Boosting; HIP, Health Independence Program; LOS, Length of Stay; OP, outpatient; IRSEAD, Index of Relative Socio-economic Advantage and Disadvantage; SHAP, Shapley Additive exPlanations; XGBoost, Extreme Gradient Boosting.

past resource use (unplanned admissions, avoidable emergency presentations and failure to attend OP appointments). Figure 3 provides a SHAP plot of each of the 18 variables.

Tables 4 and 5 present the test AUROC, sensitivity and specificity of the proposed algorithm and other models both in Australian and internationally for comparison along with the 95% CIs. Not all the referenced papers provide full details on the data sizes and performance values for their models.

DISCUSSION
The HURT algorithm had an AUROC of 84%, sensitivity of 57% and specificity of 90%. In the model, the 2% higher specificity for the HURT over the HLCC translated into 598 fewer FP and 261 more TP predictions in the 12-month time frame. The HURT algorithm also flagged

more patients that would have benefitted from targeted services who went on to have two or less unplanned admissions over 12 months.

Even though these findings are for a tertiary hospital in the state of Victoria, there are still lessons that can be applied to the broader healthcare system across Australia and internationally. In the local Australian context, the Independent Health and Aged Care Pricing Authority may apply penalties for hospitals that treat what are deemed avoidable readmissions (less than 30 days). As the HURT model has a higher specificity than other Australian models, it may be a more cost-effective tool for Australian hospitals to use as it will select less FP, and therefore, prevent hospitals who use this model from being avoidably penalised.

Researchers based in the UK have developed a number of case-finding algorithms[25–27] over the years. Direct

Table 4 Summary of test performance for predicting three or more unplanned admissions within 1 year of discharge for different case finding algorithms

Method	Country	Model	Data	Total separations (patients)	Test separations (patients)	Target: 3 or more unplanned in 12 months		
						AUROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
HURT	Australia	XGBoost	IP, ED, OP, ABS	206 714 (125 743)	34 801 (27 216)	84.2% (83.4 to 84.9)	57.2% (55.4 to 59.0)	89.4% (89.4 to 90.0)
HLCC @ NH	Australia	Weighted score	IP, ED	206 714 (125 743)	34 801 (27 216)	70.5% (69.4 to 71.8)	48.1% (46.3 to 50.0)	87.8% (87.5 to 88.2)
HLCC Victoria[17]	Australia	Weighted score	IP, ED	N/A (N/A)	N/A (N/A)	N/A (N/A)	41%* (N/A)	N/A (N/A)

*Recall was 78%.
ABS, Australian Bureau of Statistics; AUROC, Area Under the Reciever Operator Curve; ED, emergency department; HLCC, HealthLinks Chronic Care; HURT, Hospital Unplanned Readmission Tool; IP, inpatient; N/A, not available; OP, outpatient.

Method	Country	Model	Data sources	Total separations (patients)	Test separations (patients)	Target: 1 or more unplanned in 12 months		
						Auroc (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
HURT	Australia	XGBoost	IP,ED, OP, ABS	206 714 (125 743)	34 801 (27 216)	74.9% (74.2 to 75.5)	39.2% (38.2 to 40.2)	90% (89.5 to 90.4)
Billings et al[25] 2013	UK	Weighted score	IP, ED, OP, GP	N/A (1 836 099)	N/A (N/A)	78%* (N/A)	42%† (N/A)	N/A (N/A)
QAdmission Score[26]	UK	Weighted score	IP, ED, OP, GP	12 957 648 (4 870 488)	N/A (N/A)	Women: 77.3% (77.1 to 77.8) Men: 77.6% (77.4 to 77.8)	39%	N/A
SPARRA V4[27]	UK	Ensemble (ANN, RF, XGB, GLM, NB)	IP, ED, OP, GP	12 957 648 (4 870 488)	4 300 000 (N/A)	80.1% (SE: 0.023)	~52% ‡ (N/A)	~90%‡ (N/A)

Table 5 Summary of test performance predicting one or more unplanned admissions within 1 year of discharge for different case finding algorithms

*Patient-level performance.
†Recall was 78%.
‡Estimated from figure 2 (a).[27]
ABS, Australian Bureau of Statistics; ANN, Artificial Neural Network; ED, emergency department; GLM, Generalised Linear Model; GP, general practitioner; HURT, Hospital Unplanned Readmission Tool; IP, inpatient; N/A, not available; NB, Naive Bayes; OP, outpatient; RF, Random Forest; XGBoost, Extreme Gradient Boosting.

comparisons with some of these other models is difficult given data scientists use different datasets (both in terms of data captured and patient cohorts), and different definitions of an unplanned admission and benchmarks for what is considered an acceptable number of these within a 12-month period. The UK models use one or more unplanned admissions of any cause as their benchmark,[10 25–27] with the SPARRA V4 demonstrating the highest AUROC (80%) with a sensitivity of approximately 52% and a specificity of approximately 90%.[27] Where sensitivity and specificity were estimated from figure 2 (a) ROC plot.[27] Future algorithm research would benefit from application of consistent definitions so that developed algorithms may be tested and applied within different healthcare contexts (rural, remote and metropolitan) and countries.

Of particular interest in this study were the results from the SHAP scores for the importance of each feature in the HURT algorithm. Higher numbers of unplanned hospital admissions and ED admissions in the past year are shown to be important predictive features of future unplanned readmissions. In addition, lower socioeconomic status and lack of social support was predictive of unplanned readmissions, which was in agreement with SPARRA who used the Scottish Index of Multiple Deprivation using SHAP scores.[21] Both QAdmission[26] and SPARRA[27] found pathology and medication history to be an important feature for prediction of readmission, which would explain their higher performance. These data were deliberately left out so that other jurisdictions could build our model. Our next version will include this data.

The limitation of this study is that it focuses on the application of machine learning to the problem of predicting if a patient would have unplanned readmissions given current and historical information for an index admission. Hence, we only examined the performance of HURT on NH data and compared to the HLCC which was used in several Victorian health services. The model has not been subject to external validation and may not work well in non-tertiary (hospital) sites. Further work will involve multiple phases. The first phase will be to evaluate HURT within a live production system, both in terms of classification performance (eg, sensitivity) and operationally (cost savings, cohort selection). The second phase will draw on the first to improve the HURT as it is a part of a broader system that will be evaluated and optimised. Patient cohorts will be examined for FP/FN to determine any striking features that can be used or enhanced to improve ML identification of patients that will have unplanned admissions. Finally, the aim is to use general practice, pharmacy and pathology data, patient survey data and sensor in the home to better predict patients likely to readmit. These data were not included in the current approach because it is not available to the Victoria Department of Health.

CONCLUSIONS

We developed the HURT based on the XGBoost ML algorithm. We also created novel features from hospital

medical and administrative data called Cascading Chronic Conditions. The HURT algorithm was compared to the Victorian Department of Health HLCC scoring method for identifying patients at risk. The HURT model was found to have AUROC of 84%, sensitivity of 57% and specificity with 90%, 14%, 9% and 2% better than the HLCC, respectively. Future research will use pathology and pharmacy data with the aim of improving model performance.

ORCID iDs
Paul Conilione http://orcid.org/0000-0001-9913-8824
Rebecca Jessup http://orcid.org/0000-0003-2211-5231
Anthony Gust http://orcid.org/0000-0003-0509-2461

REFERENCES

1 Katterl R, Anikeeva O, Butler C, *et al*. *Potentially avoidable hospitalisations in australia: causes for hospitalisations and primary health care interventions*. 2012.
2 Spencer J. *National health and hospitals network agreement*. Australia: Council of Australian Governments, 2011.
3 van Walraven C, Bennett C, Jennings A, *et al*. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ* 2011;183:E391–402.
4 Falster MO, Jorm LR. A guide to the potentially preventable hospitalisations indicator in Australia. In: *Australian Commission on Safety and Quality in Health Care*. Sydney: Centre for Big Data Research in Health, University of New South Wales in consultation with Australian Commission on Safety and Quality in Health Care and Australian Institute of Health and Welfare, 2017: 34.
5 AIHW. Potentially preventable hospitalisations in australia by age groups and small geographic areas, 2017–18 [Overview]. 2019. Available: https://www.aihw.gov.au/reports/primary-health-care/potentially-preventable-hospitalisations/contents/about
6 OECD. *Fiscal sustainability of health systems: bridging health and finance perspectives*. 2015: 264.
7 Pang RK, Srikanth V, Snowdon DA, *et al*. Targeted care navigation to reduce hospital readmissions in "at-risk" patients. *Intern Med J* 29, 2021.
8 Artetxe A, Beristain A, Graña M. Predictive models for hospital readmission risk: a systematic review of methods. *Comput Methods Programs Biomed* 2018;164:49–64.
9 Min X, Yu B, Wang F. Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: a case study on COPD. *Sci Rep* 2019;9:2362.
10 Hung L-C, Sung S-F, Hu Y-H. A machine learning approach to predicting readmission or mortality in patients hospitalized for stroke or transient ischemic attack. *Applied Sciences* 2020;10:6337.
11 Hammoudeh A, Al-Naymat G, Ghannam I, *et al*. Predicting Hospital readmission among diabetics using deep learning. *Procedia Computer Science* 2018;141:484–9.
12 Ashfaq A, Sant'Anna A, Lingman M, *et al*. Readmission prediction using deep learning on electronic health records. *J Biomed Inform* 2019;97:103256.
13 Yang C, Delcher C, Shenkman E, *et al*. Predicting 30-day all-cause readmissions from hospital inpatient discharge data. 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom); Munich, Germany: IEEE, Munich, Germany.
14 Maali Y, Perez-Concha O, Coiera E, *et al*. Predicting 7-day, 30-day and 60-day all-cause unplanned readmission: a case study of a Sydney Hospital. *BMC Med Inform Decis Mak* 2018;18:1.
15 Jamei M, Nisnevich A, Wetchler E, *et al*. Predicting all-cause risk of 30-day Hospital readmission using artificial neural networks. *PLOS ONE* 2017;12:e0181173.
16 Division HaW. HealthLinks chronic care evaluation. In: *Summary of implementation and outcomes for 2016–17*. 1 Treasury Place, Melbourne: Victorian Government, 2019: 45.
17 Good N, Niven P, Li J, *et al*. HealthLinks: chronic care evaluation: summary report. 1 treasury place. Melbourne, Australia Victorian Government; 2022. 110.
18 Australian Bureau of Statistics. Socio-economic indexes for areas (SEIFA) 2016: commonweath government. 2016. Available: https://www.abs.gov.au/ausstats/abs@.nsf/mf/2033.0.55.001
19 Jessup RL, Bramston C, Beauchamp A, *et al*. Impact of COVID-19 on emergency department attendance in an australia hospital: a parallel convergent mixed methods study. *BMJ Open* 2021;11:e049222.
20 Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019;1:206–15.
21 Lundberg SM, Lee S-I. *A unified approach to interpreting model predictions*. Red Hook, NY, USA: Curran Associates Inc, 2017.
22 Chen T, Guestrin C. XGBoost: A scalable tree boosting system. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2016.
23 Chicco D. Ten quick tips for machine learning in computational biology. *BioData Min* 2017;10:35.
24 Kuhn M, Wing J, Weston S, *et al*. CARET: classification and regression training [internet]. 2020. Available: https://CRAN.R-project.org/package=caret
25 Billings J, Georghiou T, Blunt I, *et al*. Choosing a model to predict hospital admission: an observational study of new variants of predictive models for case finding. *BMJ Open* 2013;3:e003352.
26 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:e003482.
27 Liley J, Bohner G, Emerson SR, *et al*. Development and assessment of a machine learning tool for predicting emergency admission in scotland. *Public and Global Health* [Preprint].
28 Hart M, Dean D. *How complex is that patient*. The Health Roundtable Limited, 2006.
29 Martin C, Stockman K, Hinkley N, *et al*. Multimorbidity and acute potentially preventable diagnoses in healthlinks chronic care (HLCC) dandenong cohort. A work in evolution. *Int J Integr Care* 2021;20:173.

Supplemental Material

Table 1: Definition of features used for initial feature selection stage.

Feature	Data Type	Description
AdmissionAge	Integer	Age of patient at admission.
AdmittedOnWeekend_Flag	Binary	1 if admitted on weekend, 0 otherwise
Anxiety_drugs_Flag	Binary	1 if ICD10 diagnosis = Y47 (Drugs for anxiety), 0 otherwise
Anxiety_organic_Flag	Binary	1 if ICD10 diagnosis = F064 (Organic anxiety disorder), 0 otherwise
Anxiety_other_Flag	Binary	1 if ICD10 diagnosis = F41 (Other anxiety disorders), 0 otherwise
Anxiety_phobic_childhood_Flag	Binary	1 if ICD10 diagnosis = F931 (Phobic anxiety disorder of childhood), 0 otherwise
Anxiety_phobic_Flag	Binary	1 if ICD10 diagnosis = F40 (Phobic anxiety disorders), 0 otherwise
Anxiety_separation_childhood_Flag	Binary	1 if ICD10 diagnosis = F930 (Separation anxiety disorder of childhood), 0 otherwise
Anxiety_social_childhood_Flag	Binary	1 if ICD10 diagnosis = F932 (Social anxiety disorder of childhood), 0 otherwise
Anxiety_various_Flag	Binary	1 if ICD10 diagnosis = Y47,F41,F40,F93[0-2],F064, 0 otherwise
BirthCountry_Europe_Flag	Binary	1 if place of birth was in Europe, 0 otherwise
BirthRegion_Africa	Binary	1 if born in Africa Region, 0 otherwise
BirthRegion_Americas	Binary	1 if born in Americas Region, 0 otherwise
BirthRegion_Asia	Binary	1 if born in Asia Region, 0 otherwise
BirthRegion_Europe	Binary	1 if born in Europe Region, 0 otherwise
BirthRegion_Polar	Binary	1 if born in Polar Region, 0 otherwise
BirthRegion_Undefined	Binary	1 if born in Undefined Region, 0 otherwise
BirthSubRegion_Caribbean	Binary	1 if born in Caribbean, 0 otherwise
BirthSubRegion_Central America	Binary	1 if born in Central America, 0 otherwise
BirthSubRegion_Central Asia	Binary	1 if born in Central Asia, 0 otherwise
BirthSubRegion_Eastern Africa	Binary	1 if born in Eastern Africa, 0 otherwise
BirthSubRegion_Eastern Asia	Binary	1 if born in Eastern Asia, 0 otherwise
BirthSubRegion_Eastern Europe	Binary	1 if born in Eastern Europe, 0 otherwise
BirthSubRegion_Melanesia	Binary	1 if born in Melanesia, 0 otherwise
BirthSubRegion_Micronesia	Binary	1 if born in Micronesia, 0 otherwise
BirthSubRegion_Middle Africa	Binary	1 if born in Middle Africa, 0 otherwise
BirthSubRegion_Northern Africa	Binary	1 if born in Northern Africa, 0 otherwise
BirthSubRegion_Northern America	Binary	1 if born in Northern America, 0 otherwise
BirthSubRegion_Northern Europe	Binary	1 if born in Northern Europe, 0 otherwise
BirthSubRegion_Polar	Binary	1 if born in Polar, 0 otherwise
BirthSubRegion_Polynesia	Binary	1 if born in Polynesia, 0 otherwise
BirthSubRegion_South America	Binary	1 if born in South America, 0 otherwise
BirthSubRegion_South-Eastern Asia	Binary	1 if born in South-Eastern Asia, 0 otherwise
BirthSubRegion_Southern Africa	Binary	1 if born in Southern Africa, 0 otherwise
BirthSubRegion_Southern Asia	Binary	1 if born in Southern Asia, 0 otherwise
BirthSubRegion_Southern Europe	Binary	1 if born in Southern Europe, 0 otherwise
BirthSubRegion_Undefined	Binary	1 if born in Undefined, 0 otherwise
BirthSubRegion_Western Africa	Binary	1 if born in Western Africa, 0 otherwise
BirthSubRegion_Western Asia	Binary	1 if born in Western Asia, 0 otherwise
BirthSubRegion_Western Europe	Binary	1 if born in Western Europe, 0 otherwise
ChronicCascad_AffectiveDis_AIHW_1	Binary	1 if ICD diagnosis = (F25, F31, F34, F38, F39)
ChronicCascad_Agina_ACSC_2020	Binary	1 if ICD diagnosis = (I20, I240, I248, I249)
ChronicCascad_Anxiety_AIHW_1	Binary	1 if ICD diagnosis = (F40, F41)

Feature	Data Type	Description
ChronicCascad_Arthritis_NH_Both_1	Binary	1 if ICD diagnosis = (M13, M08, M09, U862)
ChronicCascad_Asthma_NH_Both_1	Binary	1 if ICD diagnosis = (J45, J46, J82, U833)
ChronicCascad_Bronch_NH_Both_1	Binary	1 if ICD diagnosis = (J47, U834)
ChronicCascad_CAH_NH_1	Binary	1 if ICD diagnosis = (K703, K746, K701)
ChronicCascad_Cancer_DHHSVic_1	Binary	1 if ICD diagnosis = (C, D00, D01, D02, D03, D04, D05, D06, D07, D08, D09, D37, D38, D39, D40, D41, D42, D43, D44, D45, D46, D47, D48)
ChronicCascad_CerebP_NH_Both_1	Binary	1 if ICD diagnosis = (G80, U804)
ChronicCascad_CHF_ACSC_Both_2020	Binary	1 if ICD diagnosis = (I50, J81, I110, U822)
ChronicCascad_COPD_NH_Both_1	Binary	1 if ICD diagnosis = (J20, J40, J41, J42, J43, J44, U832)
ChronicCascad_CRF_NH_Both_1	Binary	1 if ICD diagnosis = (J96, P285, U835)
ChronicCascad_Crohns_NH_Both_1	Binary	1 if ICD diagnosis = (K50, U841)
ChronicCascad_Cystic_NH_Both_1	Binary	1 if ICD diagnosis = (E84, U782)
ChronicCascad_Dement_NH_Both_1	Binary	1 if ICD diagnosis = (F00, U791)
ChronicCascad_Depr_NH_Both_1	Binary	1 if ICD diagnosis = (F0633, F32, F33, U793)
ChronicCascad_Diabetes_NH_1	Binary	1 if ICD diagnosis = (E10, E11, E13, E14)
ChronicCascad_DID_NH_Both_1	Binary	1 if ICD diagnosis = (F84, U794)
ChronicCascad_DOPA_NH_1	Binary	1 if ICD diagnosis = (K85, K86)
ChronicCascad_DOPL_NH_1	Binary	1 if ICD diagnosis = (J90, J91, J92, J93, J94) AND (C45)
ChronicCascad_DownSynd_NH_Both_1	Binary	1 if ICD diagnosis = (Q90, U882)
ChronicCascad_Emphy_NH_Both_1	Binary	1 if ICD diagnosis = (J43, U831)
ChronicCascad_Epilepsy_NH_Both_1	Binary	1 if ICD diagnosis = (G40, U803)
ChronicCascad_Gout_AIHW_1	Binary	1 if ICD diagnosis = (M10)
ChronicCascad_Hemip_NH_Both_1	Binary	1 if ICD diagnosis = (G81, U805)
ChronicCascad_HF_NH_1	Binary	1 if ICD diagnosis = (I42, I43, I50)
ChronicCascad_HypertenD_ACSC_Both_1	Binary	1 if ICD diagnosis = (I10, I119, U823)
ChronicCascad_IHD_NH_Both_1	Binary	1 if ICD diagnosis = (I20, I21, I22, I23, I24, I25, U821)
ChronicCascad_KD_NH_Both_1	Binary	1 if ICD diagnosis = (N00, N01, N02, N03, N04, N05, N06, N07, N08, N09, N10, N11, N12, N13, N14, N15, N16, N17, N18, N19, I131, I132, I139, I12, U871)
ChronicCascad_LiverFailure_NH_Both_1	Binary	1 if ICD diagnosis = (K71, U843)
ChronicCascad_Lupus_NH_Both_1	Binary	1 if ICD diagnosis = (L93, M32, U863)
ChronicCascad_MS_NH_Both_1	Binary	1 if ICD diagnosis = (G35, U802)
ChronicCascad_NIEC_NH_1	Binary	1 if ICD diagnosis = (K50, K51, K52)
ChronicCascad_Obes_NH_Both_1	Binary	1 if ICD diagnosis = (E66, U781)
ChronicCascad_Osteoa_NH_1	Binary	1 if ICD diagnosis = (M15, M16, M17, M18, M19, M47)
ChronicCascad_Osteop_Both_1	Binary	1 if ICD diagnosis = (M80, M81, M82, U864)
ChronicCascad_Parkinson_NH_1	Binary	1 if ICD diagnosis = (G20, U801)
ChronicCascad_RheuA_NH_Both_1	Binary	1 if ICD diagnosis = (M05, M06, M45, M080, U861)
ChronicCascad_RheuHD_ACSC_2020	Binary	1 if ICD diagnosis = (I00, I01, I02, I05, I06, I07, I08, I09)
ChronicCascad_Schizo_NH_Both_1	Binary	1 if ICD diagnosis = (F20, U792)
ChronicCascad_SpinaB_NH_Both_1	Binary	1 if ICD diagnosis = (Q05, U881)
ChronicCascad_Stroke_NH_1	Binary	1 if ICD diagnosis = (I60, I61, I62, I63, I64)
ChronicCascad_SubstandDis_AIHW_1	Binary	1 if ICD diagnosis = (F55, F19)
ChronicCascad_Ulcerative_NH_Both_1	Binary	1 if ICD diagnosis = (K51, U842)
Comorb_Charlson_Weighted	Integer	Weighted Charlson score (1)
Comorb_Elixhauser_Weighted	Integer	Weighted Elixhauser score (1)
ComplexityWithCCC	Integer	Number of body system treated in episode with total number of body systems impacted by cascading chronic conditions
Depression_episode_Flag	Binary	Depressive episode, 1 if icd10 diagnosis =F32, 0 otherwise

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Health Care Inform

Feature	Data Type	Description
Depression_organic_Flag	Binary	Organic depressive disorder, 1 if icd10 diagnosis =F0633, 0 otherwise
Depression_post_schizo_Flag	Binary	Post-schizophrenic depression, 1 if icd10 diagnosis =F204, 0 otherwise
Depression_recurrent_Flag	Binary	Recurrent depressive disorder, 1 if icd10 diagnosis =F33, 0 otherwise
Depression_various_Flag	Binary	Depressive conditions, 1 if icd10 diagnosis =F204,F3[2-3],F33,F0633, 0 otherwise
ExpectedLOS_LOS_var_<=-3 to <-1	Binary	1 if actual LOS of episode is between -3 and -1 standard deviations from expected LOS for given condition, 0 otherwise.
ExpectedLOS_LOS_var_<=1 to <3	Binary	1 if actual LOS of episode is between 1 and 3 standard deviations from expected LOS for given condition, 0 otherwise.
ExpectedLOS_LOS_var_<=3	Binary	1 if actual LOS of episode is less than 3 standard deviations from expected LOS for given condition, 0 otherwise.
ExpectedLOS_LOS_var_>=3	Binary	1 if actual LOS of episode is equal or greater than 3 standard deviations from expected LOS for given condition, 0 otherwise.
HAC_total	Integer	Number of Hospital Acquired Complications in episode
MaritalStatus_DevorcedWidowedSeparated_Flag	Binary	1 if divorced, widowed or separated, 0 otherwise.
NumOf_Past_Unplanned_HLCC_None_001_0to180Days	Integer	Number of clinically related unplanned admissions 180 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_001_0to30Days	Integer	Number of clinically related unplanned admissions 30 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_001_0to365Days	Integer	Number of clinically related unplanned admissions 365 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_001_0to7Days	Integer	Number of clinically related unplanned admissions 7 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_001_181to270Days	Integer	Number of clinically related unplanned admissions between 181 to 270 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_001_271to365Days	Integer	Number of clinically related unplanned admissions between 271 to 365 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_001_91to180Days	Integer	Number of clinically related unplanned admissions between 91 to 180 days prior to index admission for Chest Pain (CRG 001)
NumOf_Past_Unplanned_HLCC_None_004_0to180Days	Integer	Number of clinically related unplanned admissions 180 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_004_0to30Days	Integer	Number of clinically related unplanned admissions 30 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_004_0to365Days	Integer	Number of clinically related unplanned admissions 365 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_004_0to7Days	Integer	Number of clinically related unplanned admissions 7 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_004_181to270Days	Integer	Number of clinically related unplanned admissions between 181 to 270 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_004_271to365Days	Integer	Number of clinically related unplanned admissions between 271 to 365 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_004_91to180Days	Integer	Number of clinically related unplanned admissions between 91 to 180 days prior to index admission for Arrhythmia, Cardiac Arrest & Conduction Disorders (CRG 004)
NumOf_Past_Unplanned_HLCC_None_005_0to180Days	Integer	Number of clinically related unplanned admissions 180 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)
NumOf_Past_Unplanned_HLCC_None_005_0to30Days	Integer	Number of clinically related unplanned admissions 30 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)
NumOf_Past_Unplanned_HLCC_None_005_0to365Days	Integer	Number of clinically related unplanned admissions 365 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)
NumOf_Past_Unplanned_HLCC_None_005_0to7Days	Integer	Number of clinically related unplanned admissions 7 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)
NumOf_Past_Unplanned_HLCC_None_005_181to270Days	Integer	Number of clinically related unplanned admissions between 181 to 270 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)

Feature	Data Type	Description
NumOf_Past_Unplanned_HLCC_None_005_271to365Days	Integer	Number of clinically related unplanned admissions between 271 to 365 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)
NumOf_Past_Unplanned_HLCC_None_005_91to180Days	Integer	Number of clinically related unplanned admissions between 91 to 180 days prior to index admission for Invasive Procedure W/O AMI (CRG 005)
NumOf_Past_Unplanned_HLCC_None_006_0to180Days	Integer	Number of clinically related unplanned admissions 180 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_006_0to30Days	Integer	Number of clinically related unplanned admissions 30 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_006_0to365Days	Integer	Number of clinically related unplanned admissions 365 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_006_0to7Days	Integer	Number of clinically related unplanned admissions 7 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_006_181to270Days	Integer	Number of clinically related unplanned admissions between 181 to 270 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_006_271to365Days	Integer	Number of clinically related unplanned admissions between 271 to 365 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_006_91to180Days	Integer	Number of clinically related unplanned admissions between 91 to 180 days prior to index admission for Other Clinical Cardiology (CRG 006)
NumOf_Past_Unplanned_HLCC_None_007_0to180Days	Integer	Number of clinically related unplanned admissions 180 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_007_0to30Days	Integer	Number of clinically related unplanned admissions 30 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_007_0to365Days	Integer	Number of clinically related unplanned admissions 365 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_007_0to7Days	Integer	Number of clinically related unplanned admissions 7 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_007_181to270Days	Integer	Number of clinically related unplanned admissions between 181 to 270 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_007_271to365Days	Integer	Number of clinically related unplanned admissions between 271 to 365 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_007_91to180Days	Integer	Number of clinically related unplanned admissions between 91 to 180 days prior to index admission for Invasive Cardiac Investigative Procedures with AMI (CRG 007)
NumOf_Past_Unplanned_HLCC_None_I50_0to180Days	Integer	Number of clinically related unplanned admissions 180 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOf_Past_Unplanned_HLCC_None_I50_0to30Days	Integer	Number of clinically related unplanned admissions 30 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOf_Past_Unplanned_HLCC_None_I50_0to365Days	Integer	Number of clinically related unplanned admissions 365 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOf_Past_Unplanned_HLCC_None_I50_0to7Days	Integer	Number of clinically related unplanned admissions 7 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOf_Past_Unplanned_HLCC_None_I50_181to270Days	Integer	Number of clinically related unplanned admissions between 181 to 270 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOf_Past_Unplanned_HLCC_None_I50_271to365Days	Integer	Number of clinically related unplanned admissions between 271 to 365 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOf_Past_Unplanned_HLCC_None_I50_91to180Days	Integer	Number of clinically related unplanned admissions between 91 to 180 days prior to index admission for Invasive Cardiac Investigative Procedures with Heart Failure (I50 ICD10 3-char)
NumOfED_NonAdmitted_180Days	Integer	Calculated for past 180 days from discharge
NumOfED_NonAdmitted_270Days	Integer	Calculated for past 270 days from discharge
NumOfED_NonAdmitted_30Days	Integer	Calculated for past 30 days from discharge

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Health Care Inform

Feature	Data Type	Description
NumOfED_NonAdmitted_365Days	Integer	Calculated for past 365 days from discharge
NumOfED_NonAdmitted_3Days	Integer	Calculated for past 3 days from discharge
NumOfED_NonAdmitted_60Days	Integer	Calculated for past 60 days from discharge
NumOfED_NonAdmitted_7Days	Integer	Calculated for past 7 days from discharge
NumOfED_NonAdmitted_90Days	Integer	Calculated for past 90 days from discharge
NumOfHIPs_CHSP_12Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Commonwealth Home Support Program (CHSP) stream for past 12 months
NumOfHIPs_CHSP_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Commonwealth Home Support Program (CHSP) stream for past 6 months
NumOfHIPs_DuringSep_CHSP	Integer	Number of HIP, Commonwealth Home Support Program (CHSP) stream patient was on during unplanned episode.
NumOfHIPs_DuringSep_HACC	Integer	Number of HIP, Home and Community Care (HACC) stream patient was on during unplanned episode.
NumOfHIPs_DuringSep_HARP	Integer	Number of HIP, Hospital Admission Risk Programs (HARP) stream patient was on during unplanned episode.
NumOfHIPs_DuringSep_NDIS	Integer	Number of HIP, National Disability Insurance Scheme (NDIS) stream patient was on during unplanned episode.
NumOfHIPs_DuringSep_PAC	Integer	Number of HIP, Post Acute Care (PAC) stream patient was on during unplanned episode.
NumOfHIPs_DuringSep_ResidentialInReach	Integer	Number of HIP, Residential In Reach stream patient was on during unplanned episode
NumOfHIPs_DuringSep_SACS	Integer	Number of HIP, Sub Acute Ambulatory Care Services (SACS) stream patient was on during unplanned episode
NumOfHIPs_DuringSep_Total	Integer	Total Number of HIP patient was on during unplanned episode
NumOfHIPs_HACC_12Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Home and Community Care (HACC) stream for past 12 months
NumOfHIPs_HACC_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Home and Community Care (HACC) stream for past 6 months
NumOfHIPs_HARP_12Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Hospital Admission Risk Programs (HARP) stream for past 12 months
NumOfHIPs_HARP_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Hospital Admission Risk Programs (HARP) stream for past 6 months
NumOfHIPs_NDIS_12Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), National Disability Insurance Scheme (NDIS) stream for past 12 months
NumOfHIPs_NDIS_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), National Disability Insurance Scheme (NDIS) stream for past 6 months
NumOfHIPs_PAC_12Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Post Acute Care (PAC) stream for past 12 months
NumOfHIPs_PAC_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Post Acute Care (PAC) stream for past 6 months
NumOfHIPs_ResidentialInReach_12Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Residential Inreach Program stream for past 12 months
NumOfHIPs_ResidentialInReach_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP), Residential Inreach Program stream for past 6 months
NumOfHIPs_SACS_12Month_0Lag	Integer	Number of enrolments in any Sub Acute Ambulatory Care Services (SACS) stream for past 12 months
NumOfHIPs_SACS_6Month_0Lag	Integer	Number of enrolments in any Sub Acute Ambulatory Care Services (SACS) stream for past 6 months
NumOfHIPs_Total_12Month_0Lag	Integer	Total Number of enrolments in HIP in past 12 months
NumOfHIPs_Total_6Month_0Lag	Integer	Number of enrolments in any Hospital Independence Program (HIP) group for past 180 days
NumOfOutpatientAttended_180Days	Integer	In past 180 days from IP discharge.
NumOfOutpatientAttended_270Days	Integer	In past 270 days from IP discharge.
NumOfOutpatientAttended_30Days	Integer	In past 30 days from IP discharge.
NumOfOutpatientAttended_365Days	Integer	In past 365 days from IP discharge.
NumOfOutpatientAttended_60Days	Integer	In past 60 days from IP discharge.
NumOfOutpatientAttended_90Days	Integer	In past 90 days from IP discharge.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Health Care Inform

Feature	Data Type	Description
Outpatient_FailedToAttendRatio_180Days	Float	Calculated for past 180 days from discharge.
Outpatient_FailedToAttendRatio_270Days	Float	Calculated for past 270 days from discharge.
Outpatient_FailedToAttendRatio_30Days	Float	Calculated for past 30 days from discharge.
Outpatient_FailedToAttendRatio_365Days	Float	Calculated for past 365 days from discharge.
Outpatient_FailedToAttendRatio_60Days	Float	Calculated for past 60 days from discharge.
Outpatient_FailedToAttendRatio_90Days	Float	Calculated for past 90 days from discharge.
Outpatient_NeverBooked_180Days	Binary	No OP bookings in past 180 days.
Outpatient_NeverBooked_365Days	Binary	No OP bookings in past year.
PatientUsualAccomm_AgeCare_Alone_Flag	Binary	Living in Age Care or Alone.
Pot_Avoidable_Emergency_Admit	Binary	Patient had avoidable emergency admission according to set of ICD10 diagnosis code defined in (2)
SEIFA_Decile_Within_Australia	Integer	ABS Socio-Economic Index For Advantage, Index of Relative Socio-economic Advantage and Disadvantage (0-poor, 10-wealthy)
surgflag	Binary	Patient had surgery during episode.
TotalLOSDays_Unplanned_PastDisch_skip0days_14dayswindow	Integer	Calculated for past 14days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_180dayswindow	Integer	Calculated for past 180days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_270dayswindow	Integer	Calculated for past 270days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_30dayswindow	Integer	Calculated for past 30days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_365dayswindow	Integer	Calculated for past 365days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_3dayswindow	Integer	Calculated for past 3days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_60dayswindow	Integer	Calculated for past 60days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_7dayswindow	Integer	Calculated for past 7days from discharge.
TotalLOSDays_Unplanned_PastDisch_skip0days_90dayswindow	Integer	Calculated for past 90days from discharge.

References

1. Gasparini A. comorbidity: An R package for computing comorbidity scores. J Open Source Softw. 2018 Mar 30;3(23):648.

2. Martin C, Stockman K, Hinkley N, Campbell D. Multimorbidity and Acute Potentially Preventable Diagnoses in HealthLinks Chronic Care (HLCC) Dandenong Cohort. A work in evolution. Int J Integr Care. 2021 Feb 26;20(3):173.