Predictive validity of tools used to assess the risk of unplanned admissions: A rapid review of the evidence.

Authors

Fiona Paton, Paul Wilson, Kath Wright.

Centre for Reviews and Dissemination

University of York

York, YO10 5DD

Disclaimer

This evidence briefing presents independent research that has not been funded by an external body. The draft briefing has been produced by Fiona Paton and Paul Wilson and Kath Wright at the Centre for Reviews and Dissemination, University of York (contact: fiona.paton@york.ac.uk or paul.wilson@mbs.ac.uk).

The content of this briefing was judged to be up to date as of June 2014. The views expressed in this draft are those of the authors alone and should not be interpreted as representing the collective views of CRD research staff or the University of York.

Contribution

Fiona Paton and Paul Wilson have been involved in all stages of this rapid synthesis including production of this draft. Kath Wright designed the search strategy and performed the literature search. PMW takes overall responsibility for the draft report. He affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted.

BACKGROUND

CRD has developed an evidence briefing and support service tailored to the needs of commissioners and NHS managers (as part of the NIHR CLAHRC for Leeds York and Bradford). The service identifies, appraises and contextualises existing research evidence to inform the real world issues brought to us by local decision makers. Feedback on this service from NHS partners has been positive and we have helped some achieve major cost savings through evidence informed service reconfiguration.

The context for this particular briefing emerged from informal discussions between the evidence briefing team and John Young, the National Clinical Director for Integration and Frail & Elderly Care about evidence relating to methods and services that might be used to reduce the need for hospital admission for older people. From these discussions, a synthesis of the available evidence relating to the predictive validity of tools used to assess the risk of unplanned admissions emerged.

In 2011, the Department of Health moved from a policy of recommendation to one that embraced the plethora of prediction tools available. There are now a number of predictive tools available from commercial or academic providers and as such it was felt that a summary of their comparative performance would be of benefit.

Our aim therefore was to conduct a rapid synthesis of evidence assessing the predictive ability of tools used to identify frail elderly and people living with multiple long-term chronic health conditions who are at risk of future unplanned hospital admissions.

It should be noted that predicting the risk of future hospital readmissions are not the focus of this work. The performance of such models has already been subject to systematic review (see for example Kansagara, 2011).

METHODS

The rapid synthesis was undertaken systematically following established principles (CRD, 2009) but adapted as appropriate to ensure they are relevant to this context.

Systematic reviews, economic evaluations and other synthesised evidence (such as reviews of reviews), will be eligible for inclusion at the initial stage. If no relevant literature was identified, we then would seek to identify any primary studies evaluating predictive models that calculate the risk of future unplanned admission.

At the initial stage of the review, DARE, NHS EED and HTA databases were searched in November 2013 to identify systematic reviews and economic evaluations comparing the accuracy of different predictive risk models in identifying 'high risk' individuals who are likely to be candidates for future unplanned hospital admissions and therefore most costly.

Language and indexing used in articles identified from the initial search were used to develop searches to identify primary similar papers. The second stage of the literature screening process involved searching MEDLINE using the search strings. Reference lists of potentially relevant articles were manually searched and primary authors were contacted for additional information. Websites for relevant research centres that have published documents on predictive risk models (e.g. The Kings Fund and Nuffield Trust) were also explored.

Each stage of the screening process involved two independent researchers (FP and PMW). Any discrepancies were resolved through discussion. The inclusion criteria for the synthesis are defined as follows:

***Patients*:** Frail elderly and people living with multiple long-term chronic health conditions (eg. chronic obstructive pulmonary disease (COPD), chronic heart failure, diabetes, patients with complex needs/one or more diagnosis), at high risk of future unplanned hospital admissions and social care usage.

Exclude populations with cancer, pregnant women, and younger people with rare conditions, populations with drug and alcohol addiction or mental health conditions.

***Intervention*:** Risk screening tools used for commissioning or in primary care settings in the UK NHS or equivalent health systems to identify patients at future risk of high secondary care (health and social care) utilisation.

Studies assessing screening tools aimed at secondary care (e.g. prevention of hospital readmissions) were excluded from the review.

***Comparator**:* Reference standard (as described in the review).

Outcomes: Tools predictive ability (measured using predictive values, sensitivity and specificity, and receiver operating characteristic) to identify at risk populations to reduce unplanned hospital admissions. Clinical utility and acquisition costs.

Review processes

The intention was to assess the quality of systematic reviews and economic evaluations based on the existing critical appraisals provided by DARE and NHS EED. Other identified studies were to be appraised using criteria based on CRD guidance (CRD 2009).

Data were extracted included model description, type of source data used, populations sampled, and any indicators of predictive value, model discrimination and calibration. Data extraction was undertaken by one researcher and checked by another, Data were presented in tables and synthesised narratively to identify the tools with greatest predictive value and clinical utility.

The *c* statistic was used to describe model performance. According to Hosmer, (2000), an area under the curve (AUC) value of 0.7 to 0.8 indicates acceptable model discrimination, values of 0.8 to 0.9 indicate excellent discrimination, and values greater than 0.9 indicate outstanding discrimination. These values were used to interpret the predictive value of the models identified in the current evidence synthesis.

RESULTS

No relevant systematic reviews or economic evaluations comparing different models across primary studies were identified at the first stage.

Two hundred and eighty six references were identified at the second stage (see flow diagram on page 20). Two hundred and thirteen references were excluded on title/abstract alone. Full text articles could not be located for two references. Seventy three potentially relevant full text articles were screened. Sixty one articles were excluded as they were not relevant; 11 of which were excluded because they focused on US Medicaid or Veteran population costs, or assessed model performance based on the proportion of uncertainty explained by different models and their different variables. The purpose of these models is sometimes to set an insurance premium that is likely to cover the real cost of future healthcare (Curry 2005). One protocol was identified and the primary authors were contacted for full text articles, full publications were not identified.

Panattoni (2011) identified a number of commercial predictive risk models that can be purchased, such as 3M, Health Dialog, Verisk, D2Hawkeye, Ingenix, and MedAI (Panattoni, 2011). A separate search was run using these key terms and 18 articles were identified, only one of which was relevant to this review.

Ten articles were identified as potentially relevant to the briefing question. However, the predictive models, such as the LACE index and Patients at Risk of Readmission (PARR) aimed to predict the risk of unplanned readmissions or rates of mortality, and were therefore not relevant to the review question.

Seven studies met the inclusion criteria (Billings, 2013; Chenore, 2013; Donnan, 2008; Freund, 2013; Haas, 2013; Hippisley-Cox, 2013; Hutchings, 2013**).** The most cited models are listed in Table 1. There was an abundance of evidence relating to the use of the Adjusted Clinical Groups tool. Evidence was also available on various tools that had been adapted to local settings. For example, Chenore (2013) present evidence on the Devon Predictive Model (DPM) which was developed to explore the

factors influencing all emergency admissions in Devon (eg. age), taking into account local factors (Chenore, 2013).

The DPM includes data on inpatient stay, outpatient attendance, accident and emergency attendance from the NHS Secondary Uses Service database, combined with Devon GP data. Sixty nine variables from the Combined Predictive Model (CPM) were incorporated into a model along with local variables derived from the literature, from GPs and commissioners, and from local health data. A total of 89 variables were used to measure emergency hospital admissions in the following year (ie. unplanned hospital admissions or emergency re-admission).

The model was validated by randomly dividing the population (722,383 patients) into a derivation (training) set (80% of the population) and a validation set (using 20% of the population). The most predictive variable was age 90 to 94 years, followed by age 95 plus, and then age 85 to 89 years. Local variables that were significant predictors included shorter length of registration with the GP. At a risk score threshold of 50, the sensitivity of the DPM was 8.4%, the specificity 99.6% and the PPV 54.6%. The C-statistic was 0.781 (95% CI 0.778 to 0.783). Comparing results from the DPM and CPM showed that in each of the five highest at-risk groups, the DPM produced significantly higher PPVs than the CPM.

There was potential for coding inaccuracies in clinical records, and emergency admissions for people visiting Devon could not be captured. In addition the DPM did not separate readmissions (Chenore, 2013).

Comparisons across the different predictive models show c-statistics ranging from 0.66 (HARP simple algorithm for 30 day admission) to 0.86 (Health Numerics (RISC) model) (see Table 2). Other operational statistics such as sensitivity, specificity, and positive and negative predictive values were not consistently reported in all studies.

The HealthNumerics (RISC) model reports model performance based on three and 12 month algorithms predicting admission rates using data from acute hospital settings (c-statistic 0.855 and 0.845 respectively), and three and 12 month algorithms using data available from GP and hospital systems (0.860 and 0.852,

respectively) (personal communication). Interestingly, predictive performance reduces slightly when predicting risk of unplanned chronic admissions over a longer period of time. Similar patterns are shown for QA Admissions Score, the Hospital Admission Risk Prediction (HARP) tool, and the model explored by Billings (2013) (see Table 2).

The evidence identified tended to be based on discrimination and calibration and it was unclear whether the performance of some risk scores had been assessed in a population that was not used to develop the model (ie. had external validation). Or whether the performance of the models has been independently validated by professionals not involved in the development of the model (Hippisley-Cox, 2012). The evidence identified did not appear to formally assess the impact the models had on outcomes when used in clinical practice.

Many of the models identified were iterative. Such models incorporate additional data sets from different sources provide modest improvements in model performance. Billings (2013) assessed the added value of including different data sets for a predictive model, including data on A&E and outpatient visits, and data from GP electronic medical records. The four predictive models were: IP based on hospital inpatient data only (including day cases and regular attendances); IPAE using inpatient and A&E data; IPAEOP using inpatient, A&E and outpatient data; IPAEOPGP using inpatient, A&E, outpatient data and GP electronic medical records. The evidence is unequivocal, showing that combining risk factors (demographic, administrative, diagnostic and functional) as well as additional sources of data increases the accuracy of the model.

At the traditional risk score threshold of 50, all four models (based on data from five Primary Care Trusts) performed respectably for PPV (ranging between 0.523 for IPAEOP and 0.538 for IPAEOPGP), but sensitivity was fairly low across all four models (ranging from 0.049 for IP data only to 0.060 for the IPAEOPGP model). At a threshold of 30, the PPV reduced to between 0.417 for the IPAEOPGP model and 0.422 for the IPAE model, but sensitivity increased (ranging between 0.106 for IP only data to 0.139 for the IPAEOPGP model). The c-statistic improved with the addition of each data set, increasing from 0.731 with the inpatient-only models to

0.780 with the IPAEOPGP model. Inpatient, A&E and outpatient data can be accessed through secondary care (SUS) data but does not include social care, mental health and community data (George, 2103).

Freund (2013) compared the accuracy in hospitalisation and mortality rates of patients identified for primary care based care management as identified by primary care physicians versus predictive modelling software. Overall, primary care physicians identified 20 patients (of 464 patients per practice) as potential future care management receivers. Predictive models identified 28 patients per practice. The resources and costs required to implement care management interventions in primary care need to be taken into consideration in addition to the costs of using predictive models as case finding tools. The predictive accuracy of the models therefore needs to be superior in predicting outcomes compared to clinical knowledge (Freund 2013). In general, predictive models were shown to be superior in predicting future hospitalisation compared to physicians if comparing absolute risk of hospital admissions per patient alone, they were less efficient in taking into account the ability of patients to participate in and benefit from care management interventions. Avoidance of hospitalisations includes complex variables such as social aspects, patient behaviour which may not be captured by predictive models. Freund recommends a combined approach of predictive model selection and physician screening to complement each other, but also highlights that the proportion of patients identified by both approaches appears to be very low.

FINDINGS

There are now a large number of models available that can be used to predict the risk of unplanned hospital admissions. Overall, the models identified in this review show reasonable concordance in terms of their predictive performance (based on c-statistics). Models reporting other performance indications showed that at different thresholds, as sensitivity increased, specificity would decrease. As the algorithms become more complex or incorporate longer term horizons specificity increased but the ability of the models to identify future high cost individuals reduced. It should also be noted that whilst the reported c-statistics are broadly similar, the underlying populations, data sources and coding may differ and so this summary should not be regarded as a definitive estimate of comparative performance.

The Nuffield Trust has previously published a guide to selecting a predictive model (Lewis 2011). Factors to consider include whether to 'make or buy', the outcome to be predicted, the accuracy of the predictions made, and the availability of the data on which the model is run. Cost of the model and its software are also deemed as key. For example, whilst some included models such as the ACG and Minnesota Tiering, show reasonable predictive validity, they also require software licensing and will therefore incur license fees. Clarification of the costs (including those relating to data administration and management) associated with each model would be helpful but as yet we have not been able to ascertain estimates for those identified.

This work to date has been conducted on an unfunded basis. As such, we have not conducted exhaustive literature searches across a range of databases so there is a chance that relevant studies may have been missed. We also anticipate that there may be a substantial grey literature of potentially relevant studies that our searches and contact with providers have failed to identify or that have not been made publicly available. This phenomenon represents a form of 'publication bias' and should be borne in mind when interpreting our overall findings.

REFERENCES

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(8).

Centre for Reviews and Dissemination. *Systematic reviews: CRD's guidance for undertaking reviews in healthcare*. 3rd ed. York: Centre for Reviews and Dissemination; 2009.

Chenore T, Pereira Gray DJ, Forrer J, Wright C, Evans PH. Emergency hospital admissions for the elderly: insights from the Devon Predictive Model. *J Public Health (Oxf)* 2013; 35(4): 616-23.

Curry N, Billings J, Darin B, Curry N, Billings J, Darin B, Dixon J, Williams M, Wennberg D. *Predictive Risk Project Literature Review*. London: King's Fund, Health Dialog, New York University; 2005.

Donnan PT, Dorward DWT, 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):1416-22.

Freund T, Gondan M, Rochon J, Peters-Klimm F, Campbell S, Wensing M, et al. Comparison of physician referral and insurance claims data-based risk prediction as approaches to identify patients for care management in primary care: an observational study. *BMC Fam Pract* 2013; 14(157).

George AP, Roberts N *The use of risk stratification to understand population need and impact on health and social services. A Kent and Medway perspective*. Kent & Medway Public Health Observatory / Kent County Council, June 2013.

Georghiou T, Steventon A, Billings J, Blunt I, Lewis G, Bardsley M. *Predictive Risk and Health Care: An overview*. London: Nuffield Trust, 2011.

Haas LR, Takahashi PY, Shah ND, Stroebel RJ, Bernard ME, Finnie DM, et al. Risk-stratification methods for identifying patients for care coordination. Am J Manag Care. 2013; 19(9):725-32.

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(8).

Hosmer DG, Lemeshow S. *Applied Logistic Regression*. 2nd Ed. New York: Wiley, 2000.

Hutchings HA, Evans BA, Fitzsimmons D, Harrison J, Heaven M, Huxley P, et al. Predictive risk stratification model: a progressive cluster-randomised trial in chronic conditions management (PRISMATIC) research protocol. *Trials* 2013;14:301.

Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, Kripalani S. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011; 306 (15):1688-98.

Lewis G, Curry N, Bardsley M. *Choosing a predictive risk model: a guide for commissioners in England*. London: Nuffield Trust, 2011.

Panattoni LE, Vaithianathan R, Ashton T, Lewis GH. Predictive risk modelling in health: options for New Zealand and Australia. Aust Health Rev. 2011 Feb;35(1):45-5.

Table 1: Individual predictive risk models/risk scores

Risk Measurement Tool	Description	Social Care Data	Tool Validation Evidence Base
Adjusted Clinical Groups (ACGs), Johns Hopkins[(1)]	Around 40 CCGs in England use this system. Developed to predict the utilisation of medical resources using the presence or absence of specific diagnoses from both inpatient and outpatient services for a specified period of time, along with age and sex. Patients categorised into one of 93 discrete ACG categories with similar expected resource use into 5 Resource Utilisation Bands (RUBs). "Person-focused" approach, supporting case finding activities in the UK. **REQUIRES SOFTWARE LICENSING.** Set up and installation support available from John Hopkins staff at academic rates.	Can incorporate social care risk factors (predictive of high health care and social care costs).	Piloting of the ACG in the NHS began in 2006 in 3 PCTs. 23 page bibliography on ACG peer reviewed studies available.
CDRIntell Health Intelligence System	CDRIntell is an algorithm that can run different risk prediction models (CPM is the default model). Exports patient identifiable data from all GP clinical systems, which then validates, integrates data; population based approach for case-finding. Risk stratification built around statistical modelling, threshold modelling, clinical review, service evaluation, care planning.	NOT REPORTED	Established in 1996 in partnership with 30 CCGs; 10 hospitals and over 1,000 GP practices.
Charlson Comorbidity Index*	Sums weights for 17 conditions to predict future outcomes. **AVAILABLE BUT NEEDS TO BE PROGRAMMED TO BE APPLIED FOR LOCAL CLINICAL USE**	NOT REPORTED	Assessed in various large populations, and its validity as a prognostic measure has been demonstrated consistently.
Chronic Comorbidity Counts (CCC)	The total sum of chronic conditions is grouped into six categories: 0, 1, 2, 3, 4, and 5 or more. **AVAILABLE BUT NEEDS TO BE PROGRAMMED TO BE APPLIED FOR LOCAL CLINICAL USE**	NOT REPORTED	NOT REPORTED

Risk Measurement Tool	Description	Social Care Data	Tool Validation Evidence Base
Combined Predictive Model (CPM)	Developed to improve predictive accuracy for very high risk patients; predict risk of hospital admission for patients registered with a GP but who have not experienced a recent emergency admission; and stratify risk across all patients. Based on inpatient, outpatient, and accident & emergency data from secondary care sources as well as general practice electronic medical records. Pseudonymised risk score data available for Sussex CPM from Data Warehouse (Sussex CPM incorporates local data). Supports patient case-finding and population level analysis. **THE DEPARTMENT OF HEALTH STOPPED FUNDING THE CPM IN 2011**	NOT REPORTED	Developed with two PCTs (approx 560,000 patients) using 3 years of data; 50% used to develop the model, the other 50% to validate the model. Comparisons in the number of patients who actually had an emergency admission in the year following prediction: 586 out of top 1000 patients for the combined predictive model versus 505 out of the top 1000 for PARR.
Elder Risk Assessment Index (ERA)*	Developed to identify patients at risk for hospitalisation and emergency department visits in adults aged at least 60 years. Incorporates a weighted score of age, sex, number of hospital days in the prior two years, and marital status, as well as selected conditions (diabetes, coronary artery disease, congestive heart failure, stroke, chronic obstructive pulmonary disease, and dementia). Scores range from -1 to 34. **AVAILABLE BUT NEEDS TO BE PROGRAMMED TO BE APPLIED FOR LOCAL CLINICAL USE**	NOT REPORTED	NOT REPORTED
Hierarchical Condition Categories (HCC)	Patients are categorised into 70 conditions that contribute to a single risk score. HCCs are used to adjust Medicare capitation payments for health expenditure risk. **CAN BE DOWNLOADED**	NOT REPORTED	NOT REPORTED

Risk Measurement Tool	Description	Social Care Data	Tool Validation Evidence Base
	FROM CONTENT MANAGEMENT SYSTEM (CMS)		
Hospital Admission Risk Prediction (HARP)	Individual patient risk score of hospital admission within 30 days and 15 months. Includes data on: Patient's age; number of admissions and emergency department visits in the past six months; location where the patient was previously discharged to; intensity of previous admission; presence of the 18 top conditions; Charlson co-morbidity index; interventions during previous hospital encounter; and previous length of stay.	NOT REPORTED	Database created on a total of 385,065 initial, index episodes identified from discharges in Ontario and Manitoba (2009 to 2010). Derivation and validation models based on 191,321 acute medical episodes and 191,627 episodes, respectively using multivariate regression analysis. No completed evaluations on impact to-date.
Minnesota Tiering*	Groups patients into five 'complexity tiers' based on the number of major conditions they suffer from: low (tier 0) 0 conditions; basic (tier 1) 1 to 3 conditions; intermediate (tier 2) 4 to 6 conditions; extended (tier 3) 7 to 9 conditions; and complex (tier 4) 10 or more conditions. Currently used to determine management for care coordination among medical home plans. **REQUIRES SOFTWARE LICENSING.**	NOT REPORTED	NOT REPORTED
Person-based resource allocation formula (PBRA) (The Nuffield Trust)	Uses data from the following sources: Hospital episode statistics; Secondary uses service; NHS National Strategic Tracing Service; Quality and Outcomes Framework; and General Medical Services. Predicts individual hospital expenditure and calculates practice allocations.	NOT REPORTED	NOT REPORTED
Predict Emergency admissions Over the Next year (PEONY and PEONY II) (Scottish Executive health Department Chief Scientist Office) Tayside Centre for	Predicts risk of hospital admission for patients registered with a GP but who have not experienced a recent emergency admission. Original PEONY model fitted data reasonably well with probability of differentiating	NOT REPORTED	Based on primary care data abstracted from 40 practices throughout Scotland (n=114,421), including information on prescribing, frequency of attendance at general practices, and

Risk Measurement Tool	Description	Social Care Data	Tool Validation Evidence Base
General Practice, Community health Sciences, university of Dundee	high from low risk of 79% but with lower accuracy than the original data. Doesn't include morbidity data from primary care and the algorithm is not published or independently validated (Hippisley-Cox, 2013).		previous hospitalisation.
Predictive Risk Stratification Model (PRISM) (PRISMATIC; NHS Wales)	Estimates individual's risk of emergency admission in the following year, divided into four risk groups. Predicts risk of hospital admission for patients registered with a GP but who have not experienced a recent emergency admission. Data anonymised.	NOT REPORTED	Collates data from 37 primary care, hospital care, and demographic variables for use in primary care. No evidence from robust studies of the implementation and impact of the tool available.
QAAdmissions score (Julia Hippisley-Cox, University of Nottingham, 2013)	Estimates the risk of emergency hospital admission 1 or 2 years in patients aged 18 to 100 years in primary care. Incorporates ethnicity and clinical diagnoses, medications and abnormal laboratory results, using data solely from GP computer systems. Links to HES data using pseudonymised NHS number.	Incorporates a postcode based deprivation score, but not other information.	Based on a derivation cohort with 2,849,381 patients and validation cohort with 1,340,622 patients, which looked at demographics, lifestyle variables, chronic diseases, prescribed medication, clinical values, and laboratory test results, and number of emergency admissions in the preceding year.
Scottish Patients at risk of Readmission and Admission (SPARRA Version 3; ISD Scotland, October 2011)	Developed in 2006. Predicts an individual's risk of emergency hospital inpatient admission over the next twelve months (identifies at risk patients, using unique patient identifier, who have not experienced a recent emergency admission). Algorithm built on a linked patient-level dataset combining data from: Hospital inpatient admissions; community dispensed prescriptions; emergency department attendances; new outpatient attendances; and psychiatric inpatient admissions.	NOT REPORTED	Validation from a combined cohort consisting of approximately 3.8 million patients, separated into 3 groups; frail elderly, long term conditions, and younger emergency department.
Sussex Predictor of Key Events	The tool analyses the healthcare history of each	NOT REPORTED	An independent evaluation of the

Risk Measurement Tool	Description	Social Care Data	Tool Validation Evidence Base
(SPOKE) Sussex Health Informatics Service	resident in Sussex to predict likelihood of future admissions.		Sussex Combined Predictive Model showed that with hospital data alone the tool provides a level of predictive accuracy equivalent to national tools such as the Combined Predictive Model (CPM).
United Health UK HealthNumerics-RISC	Ensures identification of patients who can benefit from proactive care management in order to improve quality of care. Takes data from multiple sources, including primary care and acute trusts (eg. Secondary Uses Service; SUS) to perform a risk assessment of the entire population. Instructors provide training, HealthNumerics can be integrated into data warehouse.	NOT REPORTED	No published evidence, evidence based on case studies {personal communication}

* These models were compared in a retrospective cohort analysis (n=83,187 primary care patients aged at least 18 years) in the US. Limitations of the study included data restricted to provider sources only, there is therefore a risk that patients may have received care outside the study setting, and does not include outpatient pharmacy data. Study data was based on a single region with a population that is largely white and Northern European; generalisability may be limited to other populations in the US and around the world. None of the models explained more than half of the variability in outcomes, suggesting that other factors (eg. lifestyle, patient preferences) could enable better identification of patients in need of care coordination Haas (2013).

Table 2: Model performance (unplanned hospitalisations)

	C-statistic	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Adjusted Clinical Groups					
Top 5% of patients at risk	0.835{personal communication}	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
	0.73 (95% CI 0.72 to 0.73	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
Charlson Comorbidity Index					
	0.68 (95% CI 0.67 to 0.68)	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
Chronic Condition Count					
	0.69 (95% CI 0.69 to 0.70)	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
Combined Predictive Model					
	ROC 78.0%{personal communication}	NOT REPORTED	NOT REPORTED	53.8%{personal communication}	NOT REPORTED
Threshold top 1% of high risk scores	NOT REPORTED	6.0%{Georghiou, 2013}	NOT REPORTED	40.5%{Georghiou, 2013}	NOT REPORTED
Sussex Combined Predictive Model{Colin Styles, 2013}					
	ROC 81.6%	NOT REPORTED	NOT REPORTED	59.4%	NOT REPORTED
Welsh Combined Model{Georghiou, 2013}					
Threshold top 1% of high risk scores	NOT REPORTED	6.6%	NOT REPORTED	44.3%	NOT REPORTED
Devon Predictive Model(					
	0.781 (95% CI 0.778 to 0.783)	50% threshold: 8.4%	50% threshold: 99.6%	50% threshold: 54.6%	NOT REPORTED
Elder Risk Assessment (ERA) Index					
	0.71 (95% CI 0.70 to 0.72)	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
HealthNumerics (RISC)(Personal Communication)					
	Acute only (12 months): 0.845 Acute only (3 months): 0.855 Acute + GP (12 months): 0.852 Acute + GP (3 months): 0.860	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
Hierarchical Condition Categories (HCC)					
	0.67 (95% CI 0.67 to 0.68)	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED

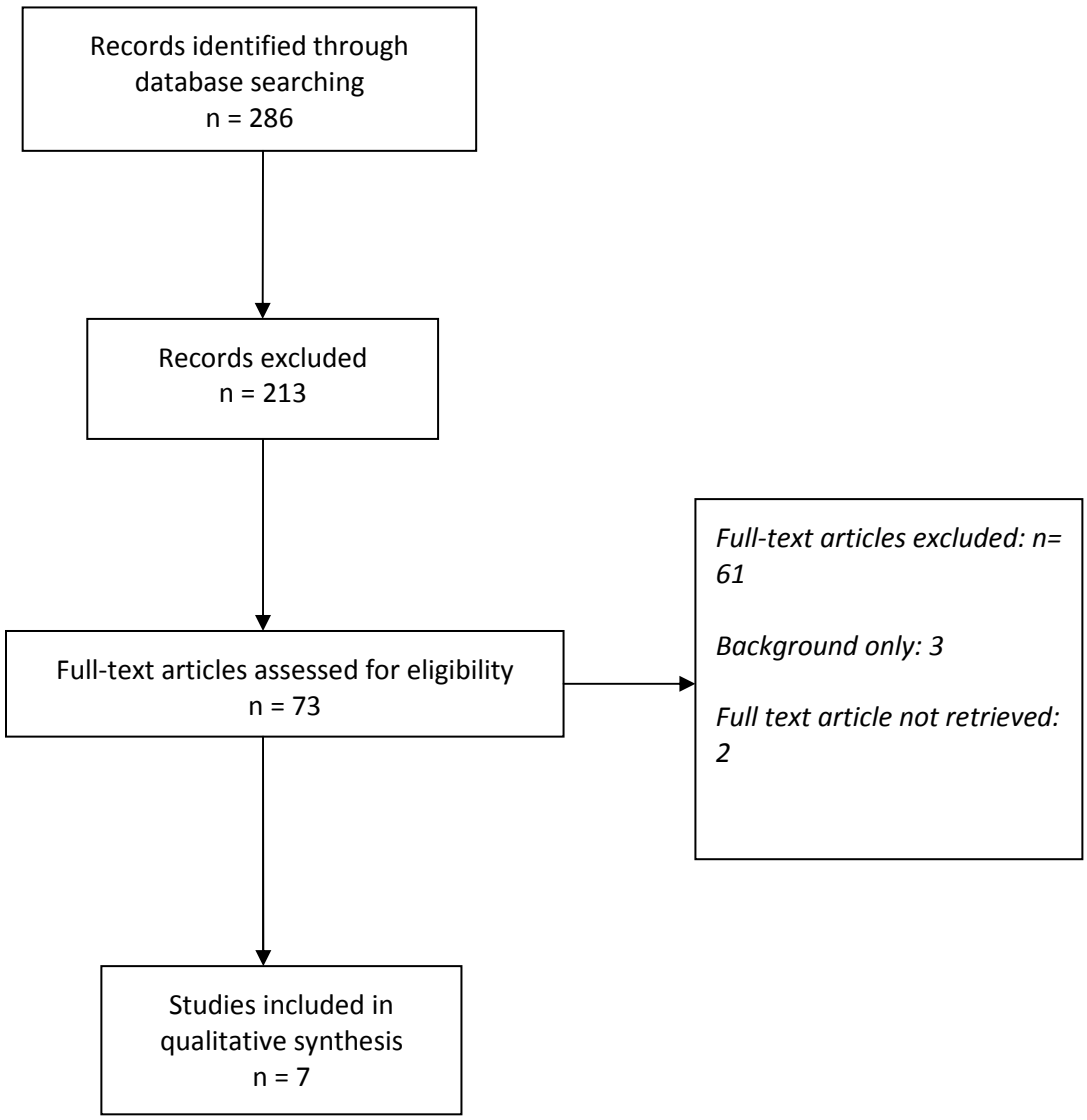
	C-statistic	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Hospital Admission Risk Prediction (HARP)					
Simple algorithm for 30 day admission	0.661	50th percentile*: 75.28% 75th percentile**: 48.82% 90th percentile***: 24.12%	50th percentile*: 45.71% 75th percentile**: 74.34% 90th percentile***: 91.39%	50th percentile*: 17.08% 75th percentile**: 22.04% 90th percentile***: 29.38% Event rate: 12.94%	50th percentile*: 92.56% 75th percentile**: 90.72% 90th percentile***: 89.02% Event rate: 87.06%
Complex algorithm for 30 day admission	0.678	50th percentile*: 74.85% 75th percentile**: 50.65% 90th percentile***: 24.26%	50th percentile*: 49.82% 75th percentile**: 74.63% 90th percentile***: 91.45%	50th percentile*: 18.14% 75th percentile**: 22.88% 90th percentile***: 29.65% Event rate: 12 .94%	50th percentile*: 93.02% 75th percentile**: 91.05% 90th percentile***: 89.04% Event rate: 87.06%
Simple algorithm for 15 month admission	0.687	50th percentile*: 67.55% 75th percentile**: 49.57% 90th percentile***: 20.19%	50th percentile*: 58.46% 75th percentile**: 75.61% 90th percentile***: 93.69%	50th percentile*: 51.48% 75th percentile**: 57.01% 90th percentile***: 67.62% Event rate: 39.49%	50th percentile*: 93.02% 75th percentile**: 91.05% 90th percentile***: 89.04% Event rate: 60.51%
Complex algorithm for 15 month admission	0.702	50th percentile*: 69.88% 75th percentile**: 42.52% 90th percentile***: 20.15%	50th percentile*: 58.59% 75th percentile**: 81.93% 90th percentile***: 93.74%	50th percentile*: 52.41% 75th percentile**: 60.55% 90th percentile***: 67.76% Event rate: 39.49%	50th percentile*: 74.88% 75th percentile**: 68.59% 90th percentile***: 64.27% Event rate: 60.51%
Minnesota Tiering					
	0.71 (95% CI 0.70 to 0.72)	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
Predict Emergency admissions Over the Next year (PEONY and PEONY II)					
PEONY	0.80	Cut-off for identifying high-risk patients (%) ≥60: 4.2% ≥49a: 7.9%	Cut-off for identifying high-risk patients (%) ≥60: 99.8% ≥49a: 99.6%	Cut-off for identifying high-risk patients (%) ≥60: 67.1% ≥49a: 59.0%	NOT REPORTED

	C-statistic	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
		≥28: 27.1% ≥18: 41.0% ≥8b: 68.9% ≥6: 76.1%	≥28: 96.8% ≥18: 92.6% ≥8b: 77.4% ≥6: 69.5%	≥28: 40.6% ≥18: 31.5% ≥8b: 19.8% ≥6: 16.8%	
PEONY{UK National Screening Committee, 2013}	NOT REPORTED	NOT REPORTED	NOT REPORTED	17% to 67% depending on risk threshold	NOT REPORTED
PEONY	NOT REPORTED	NOT REPORTED	NOT REPORTED	50% threshold: 67.1%	NOT REPORTED
QA Admissions Score					
	QResearch validation cohort (ROC) Women HES-GP linked data: 0.773 (95% CI 0.771 to 0.774) GP data alone: 0.764 (95 %CI 0.762 to 0.766) Men HES-GP linked data: 0.776 (95% CI 0.774 to 0.778) GP data alone: 0.769 (95% CI 0.767 to 0.771) CPRD validation cohort (ROC) Women HES-GP linked data: 0.771 (95% CI 0.770 to 0.773) GP data alone: 0.764 (95 %CI 0.763 to 0.766) Men HES-GP linked data: 0.772 (95% CI 0.774 to 0.778) GP data alone: 0.767 (95% CI 0.765 to 0.768)	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
Scottish Executive Health Department					

	C-statistic	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
	0.80	NOT REPORTED	NOT REPORTED	NOT REPORTED	NOT REPORTED
SPARRA					
Version 3{ISD Scotland, 2011}	NOT REPORTED	NOT REPORTED	NOT REPORTED	30% threshold: 44.1% 40% threshold: 52.2% 50% threshold: 59.8	NOT REPORTED
Versions 3{UK National Screening Committee, 2013}	NOT REPORTED	11% at 50% threshold	NOT REPORTED	60%	NOT REPORTED
Previous version{ISD Scotland, 2011}	NOT REPORTED	NOT REPORTED	NOT REPORTED	30% threshold: 46.8% 40% threshold: 53.6% 50% threshold: 59.8%	NOT REPORTED

* 1.5% of general population; **0.8% of general population; *** 0.3% of general population
ROC (received operating characteristic curve)
HES-GP (hospital episode statistics-general practitioner)
CPRD (Clinical Practice Research DataLink)

Flow Diagram
Identification of included studies



Appendix 1: Search strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:
---

1 adjusted clinical groups.ti,ab. (74)
2 ambulatory care groups.ti,ab. (27)
3 leeds risk stratificaton tool.ti,ab. (0)
4 combined predictive model.ti,ab. (5)
5 Hospital Admission Risk Profile.ti,ab. (8)
6 "Patients at risk of readmission".ti,ab. (24)
7 sparra.ti,ab. (1)
8 "Scottish Patients at Risk of Re-admission and Admission".ti,ab. (1)
9 (peony and predict$).ti,ab. (5)
10 (prism and emergency).ti,ab. (27)
11 sussex predictor.ti,ab. (0)
12 (risc and health).ti,ab. (21)
13 high impact user manager.ti,ab. (0)
14 (chads$ and health).ti,ab. (68)
15 (lace and index).ti,ab. (19)
16 medeAnalytics.ti,ab. (0)
17 hierarchical condition categories.ti,ab. (17)
18 (ccc and comorbidity).ti,ab. (2)
19 cdrIntell.ti,ab. (0)
20 health intelligence system.ti,ab. (1)
21 physician patient care alert.ti,ab. (0)
22 (PPCA and care).ti,ab. (2)
23 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (297)
24 (risk$ predict$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (1107)
25 (risk$ stratif$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (1536)
26 (risk$ adjust$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (866)
27 (risk$ screen$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (166)
28 (risk$ assess$ adj (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (2963)
29 (risk$ profil$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (81)
30 24 or 25 or 26 or 27 or 28 or 29 (6566)
31 ((predict$ and (emergency or unplanned or urgent or avoidable or preventable or unnecessary)) adj2 (hospitalisation or hospitalization or admission$ or admit$)).ti,ab. (2869)
32 23 or 30 or 31 (9611)
33 animals/ (5188900)
34 (animal or animals or monkey or monkeys or chimpanzee or chimpanzees or primate or primates or macaque or macaques or hamster or hamsters or rat or rats or mouse or mice or bird or birds or chicken or chickens or goat or goats or cattle or cow or cows or pig or pigs or dog or dogs or cat or cats or lamb or lambs or bovine or sheep or rabbit or rabbits or horse or horses or equine or camel or camels or cell or gene or cells or genes).ti. (3330754)
35 33 or 34 (6282339)
36 32 not 35 (9015)

37 (cancer$ or alcohol$ or pregnan$ or surger$ or depress$ or anxiety$ or mental$).ti. (1309066)
38 36 not 37 (7904)
39 (editorial or comment or letter).pt. (1292322)
40 38 not 39 (7769)
41 limit 40 to (english language and yr="2004 - 2014") (5553)
42 (paediatric or pediatric or child or children).ti. (534740)
43 41 not 42 (5297)

****
<76>
UI - 24328713
RO - From MEDLINE, a database of the U.S. National Library of Medicine.
ST - In-Process
AU - Sinnott JA
AU - Cai T
FA - Sinnott, Jennifer A
FA - Cai, Tianxi
IN - Department of Biostatistics, Harvard School of Public Health, 655 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.
TI - Omnibus risk assessment via accelerated failure time kernel machine modeling.
SO - Biometrics. 69(4):861-73, 2013 Dec.
AS - Biometrics. 69(4):861-73, 2013 Dec.
NJ - Biometrics
PI - Journal available in: Print-Electronic
PI - Citation processed from: Internet
JC - a5o, 0370625
OI - Source: NLM. NIHMS511697
OI - Source: NLM. PMC3869038
SB - IM
CP - United States
KW - Accelerated failure time model; Kernel machines; Omnibus test; Resampling; Risk prediction; Survival analysis
AB - Integrating genomic information with traditional clinical risk factors to improve the prediction of disease outcomes could profoundly change the practice of medicine. However, the large number of potential markers and possible complexity of the relationship between markers and disease make it difficult to construct accurate risk prediction models. Standard approaches for identifying important markers often rely on marginal associations or linearity assumptions and may not capture non-linear or interactive effects. In recent years, much work has been done to group genes into pathways and networks. Integrating such biological knowledge into statistical learning could potentially improve model interpretability and reliability. One effective approach is to employ a kernel machine (KM) framework, which can capture nonlinear effects if nonlinear kernels are used (Scholkopf and Smola, 2002; Liu et al., 2007, 2008). For survival outcomes, KM regression modeling and testing procedures have been derived under a proportional hazards (PH) assumption (Li and Luan, 2003; Cai, Tonini, and Lin, 2011). In this article, we derive testing and prediction methods for KM regression under the accelerated failure time (AFT) model, a useful alternative to the PH model. We approximate the null distribution of our test statistic using resampling procedures. When multiple kernels are of potential interest, it may be unclear in advance which kernel to use for testing and estimation. We propose a robust Omnibus Test that combines information across kernels, and an approach for selecting the best kernel for estimation. The methods are illustrated with an application in breast cancer. 2013, The International Biometric Society.
ES - 1541-0420
IL - 0006-341X

DO - http://dx.doi.org/10.1111/biom.12098
PT - Journal Article
PT - Research Support, N.I.H., Extramural
PT - Research Support, U.S. Gov't, Non-P.H.S.
NO - R01 GM079330 (United States NIGMS NIH HHS)
NO - R01 GM079330 (United States NIGMS NIH HHS)
NO - T32 CA009001 (United States NCI NIH HHS)
NO - T32 CA09001 (United States NCI NIH HHS)
NO - T32 GM074897 (United States NIGMS NIH HHS)
NO - T32 GM074897 (United States NIGMS NIH HHS)
LG - English
EP - 20131106
DP - 2013 Dec
DC - 20131216
YR - 2013
UP - 20140129

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:
---
1 adjusted clinical groups.ti,ab. (74)
2 ambulatory care groups.ti,ab. (27)
3 leeds risk stratificaton tool.ti,ab. (0)
4 combined predictive model.ti,ab. (5)
5 Hospital Admission Risk Profile.ti,ab. (8)
6 "Patients at risk of readmission".ti,ab. (24)
7 sparra.ti,ab. (1)
8 "Scottish Patients at Risk of Re-admission and Admission".ti,ab. (1)
9 (peony and predict$).ti,ab. (5)
10 (prism and emergency).ti,ab. (27)
11 sussex predictor.ti,ab. (0)
12 (risc and health).ti,ab. (21)
13 high impact user manager.ti,ab. (0)
14 (chads$ and health).ti,ab. (68)
15 (lace and index).ti,ab. (19)
16 medeAnalytics.ti,ab. (0)
17 hierarchical condition categories.ti,ab. (17)
18 (ccc and comorbidity).ti,ab. (2)
19 cdrIntell.ti,ab. (0)
20 health intelligence system.ti,ab. (1)
21 physician patient care alert.ti,ab. (0)
22 (PPCA and care).ti,ab. (2)
23 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (297)
24 (risk$ predict$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (1107)
25 (risk$ stratif$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (1536)
26 (risk$ adjust$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (866)
27 (risk$ screen$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (166)

28 (risk$ assess$ adj (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (2963)
29 (risk$ profil$ adj2 (method$ or instrument$ or model$ or index$ or score$ or tool$ or system$ or algorithm$ or rule$)).ti,ab. (81)
30 24 or 25 or 26 or 27 or 28 or 29 (6566)
31 ((predict$ and (emergency or unplanned or urgent or avoidable or preventable or unnecessary)) adj2 (hospitalisation or hospitalization or admission$ or admit$)).ti,ab. (2869)
32 23 or 30 or 31 (9611)
33 animals/ (5188900)
34 (animal or animals or monkey or monkeys or chimpanzee or chimpanzees or primate or primates or macaque or macaques or hamster or hamsters or rat or rats or mouse or mice or bird or birds or chicken or chickens or goat or goats or cattle or cow or cows or pig or pigs or dog or dogs or cat or cats or lamb or lambs or bovine or sheep or rabbit or rabbits or horse or horses or equine or camel or camels or cell or gene or cells or genes).ti. (3330754)
35 33 or 34 (6282339)
36 32 not 35 (9015)
37 (cancer$ or alcohol$ or pregnan$ or surger$ or depress$ or anxiety$ or mental$).ti. (1309066)
38 36 not 37 (7904)
39 (editorial or comment or letter).pt. (1292322)
40 38 not 39 (7769)
41 limit 40 to (english language and yr="2004 - 2014") (5553)
42 (paediatric or pediatric or child or children).ti. (534740)
43 41 not 42 (5297)

****
<76>
UI - 24328713
RO - From MEDLINE, a database of the U.S. National Library of Medicine.
ST - In-Process
AU - Sinnott JA
AU - Cai T
FA - Sinnott, Jennifer A
FA - Cai, Tianxi
IN - Department of Biostatistics, Harvard School of Public Health, 655 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.
TI - Omnibus risk assessment via accelerated failure time kernel machine modeling.
SO - Biometrics. 69(4):861-73, 2013 Dec.
AS - Biometrics. 69(4):861-73, 2013 Dec.
NJ - Biometrics
PI - Journal available in: Print-Electronic
PI - Citation processed from: Internet
JC - a5o, 0370625
OI - Source: NLM. NIHMS511697
OI - Source: NLM. PMC3869038
SB - IM
CP - United States
KW - Accelerated failure time model; Kernel machines; Omnibus test; Resampling; Risk prediction; Survival analysis
AB - Integrating genomic information with traditional clinical risk factors to improve the prediction of disease outcomes could profoundly change the practice of medicine. However, the large number of potential markers and possible complexity of the relationship between markers and disease make it difficult to construct accurate risk prediction models. Standard approaches for identifying important markers often rely on marginal associations or linearity assumptions and may not capture non-linear or interactive effects. In recent years, much

work has been done to group genes into pathways and networks. Integrating such biological knowledge into statistical learning could potentially improve model interpretability and reliability. One effective approach is to employ a kernel machine (KM) framework, which can capture nonlinear effects if nonlinear kernels are used (Scholkopf and Smola, 2002; Liu et al., 2007, 2008). For survival outcomes, KM regression modeling and testing procedures have been derived under a proportional hazards (PH) assumption (Li and Luan, 2003; Cai, Tonini, and Lin, 2011). In this article, we derive testing and prediction methods for KM regression under the accelerated failure time (AFT) model, a useful alternative to the PH model. We approximate the null distribution of our test statistic using resampling procedures. When multiple kernels are of potential interest, it may be unclear in advance which kernel to use for testing and estimation. We propose a robust Omnibus Test that combines information across kernels, and an approach for selecting the best kernel for estimation. The methods are illustrated with an application in breast cancer.
ES - 1541-0420
IL - 0006-341X
DO - http://dx.doi.org/10.1111/biom.12098
PT - Journal Article
PT - Research Support, N.I.H., Extramural
PT - Research Support, U.S. Gov't, Non-P.H.S.
NO - R01 GM079330 (United States NIGMS NIH HHS)
NO - R01 GM079330 (United States NIGMS NIH HHS)
NO - T32 CA009001 (United States NCI NIH HHS)
NO - T32 CA09001 (United States NCI NIH HHS)
NO - T32 GM074897 (United States NIGMS NIH HHS)
NO - T32 GM074897 (United States NIGMS NIH HHS)
LG - English
EP - 20131106
DP - 2013 Dec
DC - 20131216
YR - 2013
UP - 20140129