

# Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30)

John Billings,<sup>1</sup> Ian Blunt,<sup>2</sup> Adam Steventon,<sup>2</sup> Theo Georghiou,<sup>2</sup> Geraint Lewis,<sup>3</sup> Martin Bardsley<sup>2</sup>

**To cite:** Billings J, Blunt I, Steventon A, *et al.* Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Open* 2012;**00**:e001667. doi:10.1136/bmjopen-2012-001667

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2012-001667>).

Received 14 June 2012  
Accepted 29 June 2012

This final article is available for use under the terms of the Creative Commons Attribution Non-Commercial 2.0 Licence; see <http://bmjopen.bmj.com>

<sup>1</sup>Robert F. Wagner Graduate School of Public Service, New York University, New York, USA

<sup>2</sup>Nuffield Trust, London, UK

<sup>3</sup>Department of Clinical Outcomes & Analytics, Walgreen Co., Deerfield, Illinois, USA

## Correspondence to

Martin Bardsley; [martin.bardsley@nuffieldtrust.org.uk](mailto:martin.bardsley@nuffieldtrust.org.uk)

## ABSTRACT

**Objectives:** To develop an algorithm for identifying inpatients at high risk of re-admission to a National Health Service (NHS) hospital in England within 30 days of discharge using information that can either be obtained from hospital information systems or from the patient and their notes.

**Design:** Multivariate statistical analysis of routinely collected hospital episode statistics (HES) data using logistic regression to build the predictive model. The model's performance was calculated using bootstrapping.

**Setting:** HES data covering all NHS hospital admissions in England.

**Participants:** The NHS patients were admitted to hospital between April 2008 and March 2009 (10% sample of all admissions, n=576 868).

**Main outcome measures:** Area under the receiver operating characteristic curve for the algorithm, together with its positive predictive value and sensitivity for a range of risk score thresholds.

**Results:** The algorithm produces a 'risk score' ranging (0–1) for each admitted patient, and the percentage of patients with a re-admission within 30 days and the mean re-admission costs of all patients are provided for 20 risk bands. At a risk score threshold of 0.5, the positive predictive value (ie, percentage of inpatients identified as high risk who were subsequently re-admitted within 30 days) was 59.2% (95% CI 58.0% to 60.5%); representing 5.4% (95% CI 5.2% to 5.6%) of all inpatients who would be re-admitted within 30 days (sensitivity). The area under the receiver operating characteristic curve was 0.70 (95% CI 0.69 to 0.70).

**Conclusions:** We have developed a method of identifying inpatients at high risk of unplanned re-admission to NHS hospitals within 30 days of discharge. Though the models had a low sensitivity, we show how to identify subgroups of patients that contain a high proportion of patients who will be re-admitted within 30 days. Additional work is necessary to validate the model in practice.

## ARTICLE SUMMARY

### Article focus

- Preventing readmissions to hospital is important for patients, and recent policy in the English NHS means it may also impact on hospital income.
- Using logistic regression of existing person-level hospital records to develop a model that predicts the probability of readmission to hospital within 30 days.

### Key messages

- The model has been purposely designed to use only a few variables that might be entered from computerised information, or at the bedside.
- The model has reasonable accuracy in terms of positive predictive value for the highest risk patients but low sensitivity.

### Strengths and limitations of this study

- Simple and easily implemented model.
- The model has low sensitivity which means high risk patients are rare.

## INTRODUCTION

Unplanned hospital admissions and re-admissions are regarded as markers of costly, suboptimal healthcare<sup>1 2</sup> and their avoidance is currently a priority for policy-makers in many countries.<sup>3</sup> For example, in England, Department of Health guidance for the National Health Service (NHS) proposes that commissioners should not pay provider hospitals for emergency re-admission within 30 days of an index elective (planned) admission.<sup>4</sup> The rate of re-admissions will also play an important part in monitoring health system performance, as one of the new English public health 'outcome indicators'.<sup>5</sup>

In the 5-year period between 1 April 2004 and 31 March 2010, 7% of patients discharged from a hospital in England were re-admitted to hospital within 30 days,<sup>6</sup> with costs to the NHS estimated at £1.6 billion each year.<sup>7</sup> While many different interventions have been introduced with the aim of reducing unplanned admission rates,<sup>8</sup> the evidence for their efficacy and cost-effectiveness is limited.<sup>9</sup>

One reason why hospital-avoidance interventions may be unsuccessful is if they are offered to patients who are at insufficiently high risk of future unplanned hospital admission.<sup>10</sup> A history of recent hospital admissions is not an accurate predictor of future admissions by itself,<sup>11</sup> and it seems that clinicians are often unable to make reliable predictions about which patients will be re-admitted.<sup>12–13</sup> There is also some evidence to show that many re-admissions may not be avoidable.<sup>14</sup> One recent analysis observed a strong relationship between rates of rehospitalisation and overall admission rates within specific areas.<sup>15</sup>

In order to improve the accuracy of the 'case finding' process, researchers have in recent years developed a number of predictive risk models for the NHS, with the specific aim of identifying people at highest risk of a future admission or re-admission.<sup>16–21</sup> The models use relationships in routine data to identify patients at highest risk of unplanned admission or re-admission in the next 12 months. Most of these models are not contingent on an index hospital admission but instead calculate risk scores across the population at a particular date, and are designed to be run on regular (eg, monthly or quarterly) basis.

One advantage of predicting which patients are at high risk of admission in the coming 12 months is that this prolonged period may allow time for clinicians and care managers/coordinators to contact and engage with high-risk patients. Furthermore, it allows time for behavioural and treatment changes to be instigated. On the other hand, the likelihood of an unplanned admission is highest in the immediate postdischarge period,<sup>22</sup> so there may be advantages of predicting re-admissions that occur shortly after discharge. Moreover, there is evidence that some forms of preventive care may be more effective at reducing unplanned hospital admissions if initiated immediately after an acute illness.<sup>23</sup>

Outside the UK, a number of tools have been built for predicting re-admissions within 15<sup>24</sup> or 30 days<sup>25–29</sup> of discharge from hospital. Until recently, NHS funding arrangements gave hospitals in England few financial inducements to predict and prevent unplanned hospital admissions. However, the 2011–2012 operating framework proposed that NHS hospitals should not be reimbursed for re-admissions occurring within 30 days (as well as only receiving a 30% marginal rate for emergency admissions above their 2008/2009 baseline).<sup>30</sup> In practice, the degree to which this new 30-day rule is being enforced appears to vary across the country.<sup>31</sup> Yet

even without monetary incentives, knowledge of 30-day re-admission risk could still be useful to clinicians for focusing their discharge planning efforts and post-discharge support on high-risk patients.

Predictive tools built in one setting may not necessarily be accurate when used in other healthcare settings.<sup>32</sup> So in this paper, we describe how we used English hospital episode statistics (HES) data to develop a predictive model that can identify patients at high risk of re-admission to an NHS hospital in England within 30 days of discharge. The model, which we are calling 'PARR-30' (Patients at Risk of Re-admission within 30 days), can be used in practice in one of two ways: either automatically, drawing variables from Secondary Uses Service (SUS) data and from a hospital's Patient Administration System (PAS)<sup>33</sup> or 'manually' by clinicians, who can obtain the requisite information from the patient and the patient's notes and then calculate the risk using a spreadsheet or a smartphone/tablet 'app'. To facilitate this second approach, we sought to develop an algorithm that was easy to use and which relied only on a relatively small number of variables that are easily obtained from available records or from the patient. In order to justify changes in services it is often helpful to understand how the costs of the intervention may improve care and lead to lower overall costs down the line. We therefore present figures for the potential scope for savings that might accrue through reduced hospital use according to the level of risk targeted, and with assumptions about the effectiveness of interventions. We are making PARR-30 freely available for use across the NHS in England.

## METHODS

The model was developed using HES obtained from the NHS Information Centre for health and social care for the period 1 April 2006–30 April 2009.<sup>34</sup> This analysis was based on existing data that had been anonymised and therefore did not require additional ethical approval. Records were extracted for 10% of all NHS hospital admissions in England with a discharge date between 1 April 2008 and 31 March 2009. Episodes coded as births, deaths in hospital, self-discharged patients and patients transferred to other hospitals were excluded, leaving a total of 576 868 admissions remaining in the sample. Re-admissions within 30 days were restricted according to the provisions of the 2011–2012 NHS operating framework by excluding non-emergency admissions; admissions where a national tariff was not applicable; admissions for multiple trauma or transport accidents; and children aged under four. Cancer-related re-admissions were included since their exclusion in the operating framework is being reconsidered.<sup>35</sup> Patients that died after discharge were included in the development data set, reproducing what would happen if the models were applied in practice. The data set allowed

patients to have more than one re-admission episode, but each re-admission within 30 days was linked only with the most recent prior admission.

A series of logistic regressions were conducted to identify those variables that contributed most to predictions of a re-admission within 30 days of discharge, creating 'risk scores' of 0.01–1.00 describing the estimated probability of re-admission within 30 days. The variables were restricted to those that could be formulated in ways that meant they could be easily extracted from the patient or patient notes in the absence of computerised administrative data. The variables tested were based on a broad range of measures used in the PARR algorithm which predicts re-admission within the following year.<sup>14</sup> These included: the number admissions to hospital by type (emergency versus non-emergency) according to a time interval prior to current admission (90, 180, 365, 730 and 1095 days); the number of episodes per spell in prior admissions (a proxy measure of complex health problems); number of different types of specialists consulted in the last 12 months (based on services recorded in outpatient records); a range of diagnostic categories and hierarchical diagnostic groups;<sup>36</sup> characteristics of the area of residence and length of stay. A dummy variable was introduced to represent the hospital—using the largest hospital in the data as the reference point. The reduced number of variables ultimately included in this algorithm were selected based on their impact on overall model performance and ease of access to medical notes or recall by the patient.

We measured the accuracy of the predictive models in a number of ways. The positive predictive value (PPV) estimates the accuracy of the model by comparing the number of people identified by the model as being likely to experience a re-admission (based on a given threshold of risk) with the number in this group who went on to experience a re-admission. The PPV is defined as the percentage of those at-risk patients identified by the model who experience a re-admission. The sensitivity is a related concept, which measures the percentage of those people who experienced a re-admission who are correctly identified by the model as being at risk. Conversely, the specificity is defined as the proportion of people who did not experience an admission who were correctly identified as being at low risk. The sensitivity and specificity of the model can be traded off against each other by varying the threshold of risk used to define them. As well as these measures, we present estimates of the area under the receiver operating characteristic (ROC) curve, which shows the trade-off between true positives (sensitivity) and false negatives (1-specificity) at all possible thresholds. Further, we were interested in the proportion and costs of patients who experienced a re-admission by risk band (20 bands based on the level of the risk score).

Predictive models are generally 'trained' on a data set consisting of dependent variables (in this case hospital re-admissions) relating to many patients, together with a

range of independent variables from an earlier time period. The apparent performance of the model on the training (or development) data set tends to be considerably better than its performance on another, independent data set—even if that other data set consists of similar patients. In order to ensure that the model's predictions are generalisable, it is therefore important to evaluate the performance of the model more realistically than simply by calculating its accuracy on the training sample.

To do this, we used a bootstrapping evaluation method.<sup>37</sup> This method involves estimating the degree of 'optimism' associated with evaluating the apparent performance of the model on the training data set. The observed performance is moderated by subtracting the degree of optimism from the apparent performance. We calculated the degree of optimism by repeatedly drawing a large number of different bootstrapped samples from the training data set. Each consisted of the same number of patients as in the original sample, but each was formed by selecting patients randomly and allowing individual patients to be selected more than once. To estimate the optimism, we fitted models to each of these bootstrapped samples and calculated the difference between the performance of the model on the bootstrapped sample and its performance on the original sample. The optimism was estimated as the average of this quantity over all bootstrapped samples. One of the benefits of bootstrapping is that it allows all of the available patient data to be included in the data set. It has been shown to estimate model performance more accurately than other approaches such as those that involve setting aside data for a separate validation sample.<sup>38</sup>

The estimated degree of optimism was found to be very small, which we would expect given the large number of patient records available. We therefore extended the bootstrapping technique to add CI on the proportion of patients who experience a re-admission by risk band, treating optimism as negligible. These CI were formed by applying the final model to a large number (we chose 200) of bootstrapped samples, and estimating the range within which the proportions fell 95% of the time. CI were calculated for the ROC curve using a Bayesian bootstrap method.<sup>39</sup>

### Developing the business case analysis

A 'business' case analysis is presented to help guide providers and commissioners in designing interventions to prevent patient re-admissions. For this we calculated the mean re-admission costs of all patients in each risk band and at various cut-off levels. This represents the cost to NHS hospitals in terms of potential non-payment. Various assumptions are made about the effectiveness of interventions aimed at reducing the number of re-admissions within 30 days (10%, 15% and 20%), to estimate the maximum amount that could be expended on prevention, based on the estimated 'savings' from reduced admissions.

The costs of secondary care utilisation were estimated from HES data using 2010/11 Payment by Results (PbR) tariffs.<sup>40 41</sup> Activity not covered by the national tariff was costed using the national reference costs (NRC)<sup>42</sup> and adjusted to ensure that they were directly comparable with 2010/11 tariffs. If neither tariff nor NRC were available, the activity was costed as the average tariff for the specialty under which it was delivered in a method developed for a national study of resource allocation.<sup>43</sup> Therefore, costs represent income for providers rather than the actual cost of treatment for the re-admission.

We established the costs of inpatient admissions by calculating the Healthcare Resource Group (HRG) for each patient's whole stay in hospital. We derived the full cost using the PbR rules<sup>44</sup> to combine the HRG, admission method and other details of the hospital stay. This included the unit cost of the HRG and any payments due because of an unexpectedly long stay in hospital, or for any specialist care or additional treatments and tests (so-called unbundled payments). We also calculated outpatient and emergency department costs as recommended by the PbR rules.

## RESULTS

The derived model uses a small set of variable types included below:

- ▶ Patient age—used as squared value.
- ▶ Index of multiple deprivation (IMD)<sup>45</sup> for the patient's place of residence (derived from a postcode

### Box 1 A worked example of how a risk score can be calculated

An 83-year-old woman from a relatively deprived part of London is about to be discharged from a large teaching hospital in London. She received an emergency admission linked to her chronic obstructive pulmonary disease 7 days ago. Though she has not been in hospital within the last month, she did have two discharges following emergency admissions in the previous year. The patient also has a history of congestive heart failure and peripheral vascular disease.

The patient's risk of re-admission within the next 30 days was 25.1% (24.4% to 25.6%)

#### Contributions

Variable	Input	Coefficient	Term
Age squared	6889	6E-05	0.417
Number of admissions last year	2	0.121	0.243
Admission in last month	0	0.526	0.000
Current admission is emergency/unplanned	1	0.556	0.556
Deprivation—IMD score 25–40	1	0.066	0.066
Congestive heart failure	1	0.095	0.095
Peripheral vascular disease	1	0.104	0.104
Chronic pulmonary disease	1	0.224	0.224
Hospital: Barts and The London National Health Service Trust	1	0.117	0.117
Constant	1	-2.918	-2.918
Total Risk			-1.095 25.1%

**Table 1** Summary of variables\* included in model, and their coefficients, SE and significance

Variable	Coefficient	SE	Significance
Patient age (squared)	6e-5	0	<0.001
Number of emergency hospital discharges in the last year	0.121	0.002	<0.001
Whether there had been a prior emergency hospital discharge in the past 30 days	0.526	0.012	<0.001
Whether the current admission was an emergency admission	0.556	0.011	<0.001
Index of multiple deprivation band for the place of residence (lower super output area)	0.021–0.102	0.013–0.018	<0.001–0.142
History in the prior 2 years (from any hospital episode statistics primary or secondary diagnostic field) of 11 major health conditions drawn from the Charlson comorbidity index			
Congestive heart failure	0.095	0.018	<0.001
Peripheral vascular disease	0.104	0.022	<0.001
Chronic pulmonary disease	0.224	0.012	<0.001
Diabetes with chronic complications	0.146	0.032	<0.001
Renal disease	0.198	0.018	<0.001
Metastatic cancer with solid tumour	0.276	0.024	<0.001
Other malignant cancer	0.507	0.015	<0.001
Moderate/severe liver disease	0.267	0.049	<0.001
Other liver disease	0.213	0.031	<0.001
Haemiplegia or paraplegia	0.106	0.033	0.001
Dementia	0.047	0.026	0.071
Hospital-specific variable (range of values in appendix 1)	-0.976 to 0.308	0.043 to 0.206	<0.001 to 0.966
Constant	-2.918	0.032	0

\*Full details of the model and definitions are available from <http://www.nuffieldtrust.org.uk>.

**Table 2** Estimated re-admission 30 days rates and costs by risk band

Risk band	N	Total (%)	Re-admitted (%)		All patients re-admission costs		Patients with re-admission costs	
			Mean	CI	Mean	CI	Mean	CI
Band 01 (0.00–0.05)	32653	5.7	3.9	(3.6 to 4.0)	£57	(£51 to £60)	£1456	(£1366 to £1530)
Band 02 (0.05–0.10)	283165	49.1	7.1	(7.0 to 7.2)	£124	(£121 to £126)	£1747	(£1720 to £1772)
Band 03 (0.10–0.15)	146626	25.4	12.7	(12.6 to 12.9)	£298	(£293 to £306)	£2346	(£2313 to £2378)
Band 04 (0.15–0.20)	48596	8.4	18.9	(18.6 to 19.3)	£427	(£413 to £440)	£2254	(£2204 to £2313)
Band 05 (0.20–0.25)	25193	4.4	23.7	(23.2 to 24.3)	£556	(£536 to £576)	£2342	(£2276 to £2402)
Band 06 (0.25–0.30)	14282	2.5	28.0	(27.5 to 28.9)	£658	(£638 to £686)	£2347	(£2285 to £2405)
Band 07 (0.30–0.35)	8559	1.5	32.0	(31.3 to 33.0)	£765	(£733 to £802)	£2391	(£2305 to £2478)
Band 08 (0.35–0.40)	5514	1.0	36.3	(35.1 to 37.9)	£831	(£787 to £884)	£2287	(£2183 to £2370)
Band 09 (0.40–0.45)	3472	0.6	39.0	(37.4 to 41.0)	£878	(£825 to £928)	£2253	(£2140 to £2350)
Band 10 (0.45–0.50)	2413	0.4	44.9	(43.0 to 46.9)	£980	(£909 to £1051)	£2180	(£2071 to £2296)
Band 11 (0.50–0.55)	1543	0.3	47.7	(45.2 to 50.7)	£1023	(£935 to £1122)	£2143	(£2003 to £2295)
Band 12 (0.55–0.60)	1174	0.2	50.6	(48.0 to 53.3)	£988	(£916 to £1081)	£1952	(£1817 to £2094)
Band 13 (0.60–0.65)	840	0.1	54.3	(51.1 to 57.8)	£1038	(£933 to £1173)	£1912	(£1709 to £2092)
Band 14 (0.65–0.70)	617	0.1	60.6	(56.5 to 65.1)	£1148	(£1014 to £1269)	£1892	(£1716 to £2015)
Band 15 (0.70–0.75)	518	0.1	63.2	(59.8 to 67.2)	£1168	(£1041 to £1325)	£1847	(£1675 to £2054)
Band 16 (0.75–0.80)	425	0.1	65.0	(60.1 to 69.3)	£1259	(£1075 to £1423)	£1935	(£1680 to £2189)
Band 17 (0.80–0.85)	276	0.0	66.3	(60.4 to 72.4)	£1155	(£952 to £1418)	£1743	(£1444 to £2073)
Band 18 (0.85–0.90)	289	0.1	75.4	(70.2 to 80.6)	£1208	(£1037 to £1400)	£1602	(£1375 to £1803)
Band 19 (0.90–0.95)	263	0.0	83.0	(77.6 to 87.6)	£1137	(£985 to £1305)	£1369	(£1212 to £1545)
Band 20 (0.95–1.00)	450	0.1	88.7	(85.3 to 91.4)	£1189	(£1015 to £1349)	£1340	(£1137 to £1518)
All patients	576868	100	12.2	(12.1 to 12.3)	£257	(£254 to £260)	£2114	(£2098 to £2131)

Bootstrapped central estimate and 95% CI.

and mapped to one of five bands based on the lower super output area).

- ▶ Whether the current admission was an emergency admission (defined in HES as an admission category 21–28).
- ▶ Whether there had been an emergency hospital discharge in the past 30 days.
- ▶ The number of emergency hospital discharges in the last year (from any hospital).
- ▶ History in the prior 2 years (from any HES primary or secondary diagnostic field) of 11 major health conditions drawn from the Charlson co-morbidity index.<sup>46</sup>
- ▶ The hospital of the current admission, using a set of 150 dummy variables for the major acute hospitals in England.

Table 1 summarises the coefficients for these variables—the details for the individual hospital coefficients are provided in appendix 1. Box 1 gives an example of how a risk score for an individual patient could be calculated. Full details of the model will also be made available on the Nuffield Trust website (<http://www.nuffieldtrust.org.uk>).

The performance of the model is shown in table 2 in terms of the percentage of patients with a 30-day re-admission, and the costs of those re-admissions displayed by risk band vingtiles. For the higher-risk patients (risk bands 11 and above), re-admission rates ranged from 47.7% to 88.7% compared with an overall re-admissions rate of 12.2%. However, the number of patients in these high-risk bands represented only a

small share (1.1%) of all patients analysed. For risk bands 1–10, the risk of re-admission within 30 days dropped steadily with decreasing risk score, but the number of patients in each band increased. The two lowest-risk bands cover 54.7% of patients with a risk of re-admission within 30 days of 7.1% or lower.

The mean re-admission costs tended to be lower in the lower-risk bands because a smaller percentage of patients were re-admitted. However, those in the lower bands who had a re-admission tended to have higher costs (eg, £1340 per admission for patients in band 20 compared with £2143 per admission for patients in band 11).

A business case analysis is provided in table 3, documenting the rate of re-admissions and the maximum level of expenditure at each risk band (and at various risk band cut-off levels). These values indicate where the cost of the preventive intervention equals the net savings from reduced re-admissions—with various assumptions about the effectiveness of interventions (10%, 15% and 20%). With a risk band cut-off at Band 11, mean re-admission costs were £1088 (CI £1046 to £1124—not shown) per patient. Using an assumption of a 10% reduction in the rate of re-admission, £109 per patient (CI £105 to £112—not shown) could be spent on the 6395 patients in these bands, with the costs of the intervention equalling costs of avoided emergency admissions (breakeven).

The PPV for the model for all patients with a risk score above 0.50 (risk bands 11+) was 59.2% (CI 58.0% to 60.5%), with specificity of 99.5% (CI 99.5% to 99.5%) and sensitivity of 5.4% (CI 5.2% to 5.6%; see table 4).

**Table 3** 'Business Case' analysis

Risk band	N	By risk band level			N	Cumulative at band cut-off level				Maximum expenditure per patient for break even*		
		Maximum expenditure per patient for break even				All patients (%)	Pats with re-adm (%)*	All pats with re-adm (%)*	Mean re-adm cost*	Assumed reduction in Re-admission		
		10%	15%	20%						10%	15%	20%
Band 01 (0.00–0.05)	32653	£6	£9	£11	576868	100	12.2	100	£257	£26	£39	£51
Band 02 (0.05–0.10)	283165	£12	£19	£25	544215	94.3	12.7	98.2	£269	£27	£40	£54
Band 03 (0.10–0.15)	146626	£30	£45	£60	261050	45.3	18.7	69.5	£427	£43	£64	£85
Band 04 (0.15–0.20)	48596	£43	£64	£85	114424	19.8	26.3	42.9	£591	£59	£89	£118
Band 05 (0.20–0.25)	25193	£56	£83	£111	65828	11.4	31.8	29.8	£713	£71	£107	£143
Band 06 (0.25–0.30)	14282	£66	£99	£132	40635	7.0	36.8	21.3	£809	£81	£121	£162
Band 07 (0.30–0.35)	8559	£76	£115	£153	26353	4.6	41.6	15.6	£892	£89	£134	£178
Band 08 (0.35–0.40)	5514	£83	£125	£166	17794	3.1	46.2	11.7	£953	£95	£143	£191
Band 09 (0.40–0.45)	3472	£88	£132	£176	12280	2.1	50.7	8.9	£1008	£101	£151	£202
Band 10 (0.45–0.50)	2413	£98	£147	£196	8808	1.5	55.3	6.9	£1059	£106	£159	£212
Band 11 (0.50–0.55)	1543	£102	£153	£205	6395	1.1	59.2	5.4	£1088	£109	£163	£218
Band 12 (0.55–0.60)	1174	£99	£148	£198	4852	0.8	62.8	4.3	£1109	£111	£166	£222
Band 13 (0.60–0.65)	840	£104	£156	£208	3678	0.6	66.7	3.5	£1148	£115	£172	£230
Band 14 (0.65–0.70)	617	£115	£172	£230	2838	0.5	70.3	2.8	£1180	£118	£177	£236
Band 15 (0.70–0.75)	518	£117	£175	£234	2221	0.4	73.0	2.3	£1189	£119	£178	£238
Band 16 (0.75–0.80)	425	£126	£189	£252	1703	0.3	76.0	1.8	£1196	£120	£179	£239
Band 17 (0.80–0.85)	276	£115	£173	£231	1278	0.2	79.7	1.5	£1175	£118	£176	£235
Band 18 (0.85–0.90)	289	£121	£181	£242	1002	0.2	83.4	1.2	£1181	£118	£177	£236
Band 19 (0.90–0.95)	263	£114	£171	£227	713	0.1	86.6	0.9	£1170	£117	£175	£234
Band 20 (0.95–1.00)	450	£119	£178	£238	450	0.1	88.7	0.6	£1189	£119	£178	£238
All patients	576868	£26	£39	£51								

Estimates of potential savings that could be made to fund an intervention, achieved at different risk bands and with differing assumptions about the reduction in admissions achieved.

\*CI and other details on the model are available at <http://www.nuffieldtrust.org.uk/our-work/projects/predicting-risk-hospital-readmission-parr-30>.

**Table 4** Estimated model performance bootstrapped central estimate and 95% CI

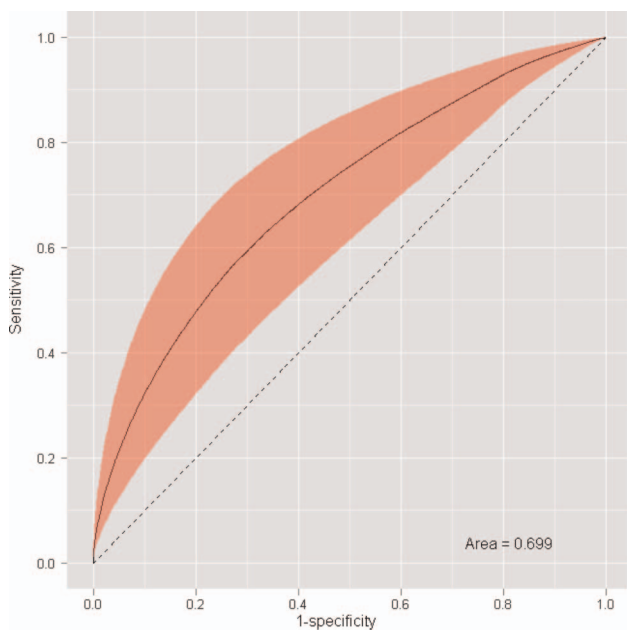
	Central estimate (%)	CI
Positive predictive value*	59.2	(58.0 to 60.5%)
Sensitivity*	5.4	(5.2 to 5.6%)
Specificity*	99.5	(99.5 to 99.5%)
Area under the receiver operating characteristic curve	0.70	(0.69 to 0.70)

\*Data are for a risk score threshold 50+.

The ROC curve ROC in figure 1 illustrates the trade-off between true positives (sensitivity) and false negatives (1—specificity) for the model. Overall, the area under the curve was 0.70 (CI 0.69 to 0.70).

### DISCUSSION

We have built a predictive model using a limited set of variables that were generated from HES. The model estimates the risk and costs of re-admission to an NHS hospital in England within 30 days of discharge. We have intentionally selected variables that we believe will easily translate to information available from patients' notes or from the patients themselves. Look-up tables can be built to map variables such as a patient's postcode to deprivation score. This means it is possible to build simple software tools such as a spreadsheet or 'app' to calculate scores, as well as by using data from a hospital's patient administration system.



**Figure 1** Receiver operating characteristic curve for the bootstrapped central estimate (red line) and 95% CI (shaded area).

The performance of the model was respectable, with a PPV of 59.2% for a risk score threshold of 50+ and an area under the ROC curve ('c-statistic') of 0.70. For example, a recent systematic review of predictive risk models for 30 day re-admissions documented c-statistics ranging from 0.50 to 0.72.<sup>47</sup> The specificity of this model (99.5%) is high, although the sensitivity of the model is quite low with only 5.4% of all patients in the sample (bands 11+). The performance of the model could have been improved by including more variables but this would have made the model less useful in practice. Traditional measures of performance, such as the sensitivity, mask the potential value of models in targeting preventive interventions. Knowledge of the percentage of patients in each risk score band who will have an admission in the next 30 days can be useful in titrating resources to patients, with more or different types of resources assigned for patients who are most likely to have a hospital admission. At the highest-risk band, patients had an 88.7% chance of hospital re-admission within 30 days and £178 could be spent per patient on interventions aimed at avoiding re-admission, assuming these interventions were successful at averting 15% of all re-admissions and that breakeven was required. The level and type of resources allocated to these patients should be different from those allocated to patients in the lower-risk levels, such as those in band 6 where chances of re-admission were 28.0%. These data can also be used in setting an overall cut-off level/threshold for the full range of intervention strategies. For example, at a cut-off level at band 5, almost 30% of patients who will have an admission in the next 30 days will be included, and the chance of these patients having a re-admission is 31.8%. The levels and type of intervention for these patients should vary by risk band and patient characteristics, but clinicians and commissioners can use these data to select thresholds for any preventive intervention.

The model has its limitations. It was developed using HES data, but it is intended to be used by hospitals using either a combination of PAS data and SUS data or patient self-reported information on prior use and medical history from the patient's notes. While PAS/SUS data do differ from HES, the differences are minor so we believe this shortcoming is unlikely to affect the accuracy of the predictive model substantially. However, differences in patients' recall of their prior hospital use and their medical history present bigger challenges to the validity of the model. Self-recall data on health care utilisation can differ from administrative data, especially for people with high levels of healthcare use, older people and people with poor health status.<sup>48 49</sup> We are currently testing the model to determine the extent to which patient-reported information differs from that recorded in HES.

The ability to identify patients at high risk of re-admission constitutes the first step in any strategy to improve care and services for susceptible patients. The ultimate goal, however, is to couple this 'case finding'

process with cost-effective interventions that mitigate the risk of re-admission, and ideally, uses the ensuing financial savings to help fund the intervention. Unfortunately, only a modest amount is known about what works, and for whom, in reducing re-admissions.

In a recent systematic review, Hansen *et al*<sup>50</sup> identified a broad range of strategies that have been employed, including predischarge interventions (improved discharge planning, patient education, medication reconciliation, postdischarge follow-up appointment, etc), postdischarge interventions (patient hotlines, telephone appointment reminders, home visits, etc) and other interventions to bridge the transition from hospital to home such as nurse coaching. Many of the studies looked at were small and not well designed. Five of 16 randomised controlled trials documented statistically significant reductions in the absolute risk of re-admission, but no single intervention or bundle of strategies were found to be consistently successful in reducing risk.

The data on costs developed here also suggest additional caution. At a risk score cut-off of 0.50 (band 11+), even with an optimistic assumption of a 20% reduction in the rate of re-admissions, the amount available to spend on an intervention and still achieve breakeven is relatively modest (£218 per patient). Broadening the intervention to a cut-off at band 5, this amount drops to £143 (and £71 if a more realistic reduction in re-admissions of 10% is assumed), see table 3. While improved discharge planning, arranging postdischarge follow-up visits and telephone reminders may be relatively inexpensive, other interventions such as nurse coaching and home visits can become quite costly. These data would permit targeting of interventions, with more costly strategies limited to the patients at highest risk, but the level of available resource will undoubtedly be strained if breakeven is expected.

As hospitals in England begin responding to the new financial incentives included in the 2011–2012 operating framework, it will be important to gather evidence about what interventions are effective and for which patients and at what cost. Areas for future research may include determining whether and how the effectiveness of interventions differs according to the underlying level of risk. For example, it may be that patients at lower or moderate risk of re-admission have conditions or circumstances where an intervention is more likely to succeed than for patients at high risk. Equally, there may be certain subgroups of patients within a particular risk band who are more or less amenable to preventive care. The use of predictive models as case finding tools to target preventive interventions has gained considerable currency in community-based settings. We believe that it is important to consider how such tools might be used in the much more immediate care environment of the hospital to improve the long-term management of patients.

**Acknowledgements** We are grateful to the staff at Chelsea & Westminster Hospital and the Royal Berkshire Hospital for their support and guidance as we developed the PARR-30 model.

**Contributors** The preparation of data sets and input variables and costs were undertaken by TG and IB; J B did the central modelling and reporting while AS undertook work on bootstrapping and testing derived models. GL wrote the first draft of the paper and coordinated advice from local sites. MB advised on the analysis and results and managed the work of the research team at the Nuffield Trust. All authors contributed to the writing of the paper. GL was employed as a Senior Fellow at the Nuffield Trust at the time this work was undertaken. JB is the guarantor.

**Funding and disclaimer** This research was funded by the Nuffield Trust. The study sponsor was the Chairman of the Nuffield Trust. The sponsor had no role in and the collection, analysis and interpretation of data, in the writing of the article nor in the decision to submit it for publication.

**Competing interests** All authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare that no authors have any relationships with any companies that might have an interest in the submitted work in the previous 3 years; none of their spouses, partners or children have any financial relationships that may be relevant to the submitted work; and no authors have any non-financial interests that may be relevant to the submitted work.

**Ethical approval** This study only involved the analysis of pseudonymous secondary data. Since there were no identifiable human subjects, ethics approval was not required for this research and informed consent was not sought.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** Details of the derived models variables and definition are available from the authors at the nuffield trust at [research@nuffieldtrust.org.uk](mailto:research@nuffieldtrust.org.uk).

## REFERENCES

- McKee M, Nolte E. Chronic care. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S *Performance measurement for health system improvement: experiences, challenges and prospects*. Cambridge: Cambridge University Press, 2009.
- Billings J, Zeitel L, Lukomnik J, *et al*. Impact of socioeconomic status on hospital use in New York City. *Health Aff* 1993; 12:162–73.
- Health System Priorities in the Aftermath of the Crisis. Paris: OECD, 2010. <http://www.oecd.org/dataoecd/14/36/46098360.pdf> (accessed 10 Sept 2011).
- Department of Health. *Payment by Results Guidance for 2011–12*. London: Department of Health, 2011.
- Department of Health. *Healthy Lives, Healthy People: transparency in outcomes, proposals for a public health outcomes framework*. London: Department of Health, 2010.
- Blunt I, Bardsley M, Clarke A. Analysis of emergency 30-day readmissions in England using routine hospital data 2004–2010. Is there scope for reduction? (Forthcoming)
- Robinson P. Hospitals readmissions and the 30 day threshold. London: CHKS, 2010 [http://www.chks.co.uk/assets/files/Hospital\\_readmissions\\_and\\_the\\_30\\_day\\_threshold\\_final.pdf](http://www.chks.co.uk/assets/files/Hospital_readmissions_and_the_30_day_threshold_final.pdf) (accessed 2 Dec 2011).
- Purdy S. Avoiding hospital admissions: what does the research evidence say? London: King's Fund, 2011. [http://www.kingsfund.org.uk/publications/avoiding\\_hospital.html](http://www.kingsfund.org.uk/publications/avoiding_hospital.html) (accessed 10 Sept 2011).
- Hansen LO, Young RS, Hinami K, *et al*. Interventions to reduce 30-day rehospitalization: a systematic review. *Ann Intern Med* 2011; 155:520–8.
- Steventon A, Bardsley M, Billings J, *et al*. An evaluation of the impact of community-based interventions on hospital use. London: Nuffield Trust, 2011. [http://www.nuffieldtrust.org.uk/sites/files/nuffield/an-evaluation-of-the-impact-of-community-based-interventions-on-hospital-use-full\\_report.pdf](http://www.nuffieldtrust.org.uk/sites/files/nuffield/an-evaluation-of-the-impact-of-community-based-interventions-on-hospital-use-full_report.pdf) (accessed 11 Sept 2011).
- Roland M, Dusheiko M, Gravelle H, *et al*. Follow up of people aged 65 and over with a history of emergency admissions: analysis of routine admission data. *BMJ* 2005; 330:289–92.
- Curry N, Billings J, Darin B, *et al*. Predictive risk Project Literature Review. London: King's Fund, 2005. <http://www.kingsfund.org.uk/document.rm?id=6196> (accessed 12 June 2012).
- Allaudeen N, Schnipper JL, Orav EJ, *et al*. Inability of providers to predict unplanned readmissions. *J Gen Intern Med* 2011; 26:771–6.



14. van Walraven C, Bennett C, Jennings A, *et al*. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ* 2011;183:E391–402.
15. Epstein AM, Jha AK, Orav EJ. The relationship between hospital admission rates and rehospitalizations. *N Engl J Med* 2011;365:2287–95.
16. Billings J, Dixon J, Mijanovich T, *et al*. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *Br Med J* 2006;333:327.
17. Wennberg D, Siegel M, Darin B, *et al*. Combined predictive model: final report and technical documentation. London: Health Dialog/King's Fund/New York University, 2006. <http://www.kingsfund.org.uk/document.rm?id=8248> (accessed 10 Sept 2011).
18. SPARRA: Scottish Patients At Risk of Readmission and Admission. Edinburgh: NHS National Services Scotland, 2006. [http://www.isdscotland.org/isd/files/SPARRA\\_Report.pdf](http://www.isdscotland.org/isd/files/SPARRA_Report.pdf) (accessed 10 Sept 2011).
19. NHS Scotland Information Services Division. *SPARRA Mental Health: Scottish Patients at Risk of Readmission and Admission (to psychiatric hospitals or units)*. Edinburgh: NHS Scotland Information Services Division, 2009. <http://www.isdscotland.org/Health-Topics/Health-and-Social-Community-Care/SPARRA/SPARRA-History/SPARRA-MH-report-final.doc> (accessed 10 Sept 2011).
20. Donnan PT, Dorward DW, Mutch B, *et al*. Development and validation of a model for predicting emergency admissions over the next year (PEONY): a UK historical cohort study. *Arch Intern Med* 2008;168:1416–22.
21. NHS Wales Informatics Service. *Reducing emergency risk: the prism tool*. NHS Wales Informatics Service. <http://www.wales.nhs.uk/nwis/page/52558> (accessed 12 June 2012).
22. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the medicare fee-for-service program. *N Engl J Med* 2009;360:1418–28.
23. Puhan MA, Scharplatz M, Troosters T, *et al*. Respiratory rehabilitation after acute exacerbation of COPD may reduce risk for readmission and mortality: a systematic review. *Respir Res* 2005;6:54.
24. Rowland K, Maitra AK, Richardson DA, *et al*. The discharge of elderly patients from an accident and emergency department: functional changes and risk of readmission. *Age Ageing* 1990;19:415–18.
25. Hasan O, Meltzer DO, Shaykevich SA, *et al*. Hospital readmission in general medicine patients: a prediction model. *J Gen Intern Med* 2010;25:211–19.
26. Novotny NL, Anderson MA. Prediction of early readmission in medical inpatients using the probability of repeated admission instrument. *Nurs Res* 2008;57:406–15.
27. van Walraven C, Dhalla IA, Bell C, *et al*. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *CMAJ* 2010;182:551–7.
28. Salvi F, Morichi V, Grilli A, *et al*. Predictive validity of the Identification of Seniors At Risk (ISAR) screening tool in elderly patients presenting to two Italian Emergency Departments. *Ageing Clin Exp Res* 2009;21:69–75.
29. Meldon SW, Mion LC, Palmer RM, *et al*. A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department. *Acad Emerg Med* 2003;10:224–32.
30. Department of Health. *The operating framework for the NHS in England 2011/12*. London: Department of Health, 2010.
31. Dowler C. Emergency readmissions payment ban being ignored. *Health Serv J* 2011. <http://www.hsj.co.uk/news/finance/emergency-readmissions-payment-ban-being-ignored/5032666>. article (accessed 11 Sept 2011).
32. Fan J, Worster A, Fernandes CM. Predictive validity of the triage risk screening tool for elderly patients in a Canadian emergency department. *Am J Emerg Med* 2006;24:540–4.
33. NHS Information Centre for Health and Social Care. *The SUS programme*. Leeds: NHS Information Centre for Health and Social Care, 2011. <http://www.connectingforhealth.nhs.uk/systemsandservices/sus> (accessed 12 Sept 2011).
34. Information Centre for Health and Social Care. *Hospital episode statistics*. Leeds: Information Centre, 2009.
35. Payment by Results Guidance for 2011–12. London: Department of Health, 2011. [http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/documents/digitalasset/dh\\_126157.pdf](http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_126157.pdf) (accessed 12 June 2012).
36. Pope GC, Kautter J, Ellis RP, *et al*. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev* 2004;25:119–41.
37. Efron B, Tibshirani R. An introduction to the bootstrap. In: *Monographs on statistics and applied probability*. New York: Chapman and Hall; 1993; p. xvi, 436.
38. Steyerberg EW, Harrell FE Jr, Borsboom GJJM, *et al*. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.
39. Gu J, Ghosal S, Roy A. Bayesian bootstrap estimation of ROC curve. *Stat Med* 2008;27:5407–20.
40. Department of Health. *Payment by results guidance for 2010–11*. London: Department of Health, 2010.
41. PBRA3 team. Updating and enhancing a resource allocation formula at general practice level based on individual level characteristics (person-based resource allocation). (Forthcoming)
42. Department of Health. *NHS reference costs 2007–08*. London: Department of Health, 2009.
43. Dixon J, Smith P, Gravelle H, *et al*. A person based formula for allocating commissioning funds to general practices in England: development of a statistical model. *BMJ* 2011;343:d6608. doi: 10.1136/bmj.d6608 (Published 22 November 2011).
44. Department of Health. *Payment by results guidance for 2010–11*. London: Department of Health, 2010b.
45. Department for Communities and Local Government. *The English Indices of Deprivation 2010*. <http://www.communities.gov.uk/publications/corporate/statistics/indices2010> (accessed 12 June 2012).
46. Charlson ME, Pompei P, Ales KL, *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
47. Kansagara D, Englander H, Salanitro A, *et al*. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;306:1688–98.
48. Bellón JA, Lardelli P, de Dios LJ, *et al*. Validity of self reported utilisation of primary health care services in an urban population in Spain. *J Epidemiol Community Health* 2000;54:544–51.
49. Cleary PD, Jette AM. The validity of self-reported physician utilization measures. *Med Care* 1984;22:796–803.
50. Hansen LO, Young RS, Hinami K, *et al*. Interventions to reduce 30-day rehospitalisation: a systematic review. *Ann Intern Med* 2011;155:520–8.