

BMJ Open Observational study to determine the utility of hospital administrative data to support case finding of English patients at higher risk of severe healthcare-related harm

Helen Hogan,¹ Nora Cooke-O'Dowd,² Kaushik Chattopadhyay,¹ Jan van der Meulen,¹ Christopher Sherlaw-Johnson,² Nick Black¹

To cite: Hogan H, Cooke-O'Dowd N, Chattopadhyay K, *et al*. Observational study to determine the utility of hospital administrative data to support case finding of English patients at higher risk of severe healthcare-related harm. *BMJ Open* 2019;**9**:e025372. doi:10.1136/bmjopen-2018-025372

► 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-2018-025372>).

Received 16 July 2018
Revised 20 May 2019
Accepted 23 May 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Health Services Research and Policy, London School of Hygiene & Tropical Medicine, London, UK
²Nuffield Trust, London, UK

Correspondence to

Dr Helen Hogan;
helen.hogan@lshtm.ac.uk

ABSTRACT

Objectives To identify ways of using routine hospital data to improve the efficiency of retrospective reviews of case records for identifying avoidable severe harm

Design Development and testing of thresholds and criteria for two indirect indicators of healthcare-related harm (long length of stay (LOS) and emergency readmission) to determine the yield of specified harms coded in Hospital Episode Statistics (HES).

Setting Acute National Health Service hospitals in England.

Participants HES for acute myocardial infarction (AMI), bowel cancer surgery and hip replacement admissions from 2014 to 2015.

Interventions Case-mix-adjusted linear regression models were used to determine expected LOS. Different thresholds were examined to determine the association with harm. Screening criteria for readmission included time to readmission, length of readmission and diagnoses in initial admission and readmission. The association with harm was examined for each criterion.

Results The proportions of AMI cases with a harm code increased from 14% among all cases to 47% if a threshold of three times the expected LOS was used. For hip replacement the respective increase was from 10% to 51%. However as the number of patients at these higher thresholds was small, the overall proportion of harm identified is relatively small (15%, 19%, 9% and 8% among AMI, urgent bowel surgery, elective bowel surgery and hip replacement cohorts, respectively). Selection of the time to readmission had an effect on the yield of harms but this varied with condition. At least 50% of surgical patients had a harm code if readmitted within 7 days compared with 21% of patients with AMI.

Conclusions Our approach would select a substantial number of patients for case record review. Many of these cases would contain no evidence of healthcare-related harm. In practice, Trusts may choose how many reviews it is feasible to do in advance and then select random samples of cases that satisfy the screening criteria.

Strengths and limitations of the study

- Routine hospital administrative data is inexpensive and easy to access.
- Potential healthcare-related harm can be identified in these data by specific codes used for such harms for example, complications and adverse reactions or by using indirect indicators known to be linked to such harm such as long length of stay (LOS) or readmission.
- Comparing the performance of long LOS and readmission across four contrasting cohorts of patients (emergency: acute myocardial infarction; urgent: non-elective bowel cancer surgery; semiurgent: elective bowel cancer surgery; and elective: hip replacement) when thresholds for LOS and criteria for readmission are manipulated shows that sensitivity and positive predictive power to identify harm can be increased.
- To confirm if any harm identified in the administrative records is healthcare-related, retrospective case record review is required.
- The approach would identify potential healthcare-related harm in large numbers of cases. A selection process for those going forward to case record review would be required.

INTRODUCTION

There are two main ways avoidable severe harm is identified in patients in acute hospitals in the National Health Service (NHS) in England. However, both approaches have shortcomings. Incident reporting systems depend on staff compliance while retrospective case record reviews (RCRR) requires considerable resources which preclude universal application to all hospital admissions.

An alternative approach could be for hospitals to employ administrative data to screen records for case note review. These include

codes for healthcare-related harm ('direct' indicators of harm) such as 'complications and adverse events' or 'pulmonary embolism after a surgical procedure'. In theory, all harm should be recorded though the completeness of recording is doubtful. In addition, the data will not distinguish between levels of severity so instances of severe harm cannot be distinguished from lesser forms. There has been no recent estimation of the completeness of reporting of harm in administrative data, but a historical (1999–2003) comparison with Australia suggested under-reporting: 2.2% of all NHS admissions compared with 4.75% in Australia.¹²

Administrative data also offers the possibility of using two 'indirect' indicators that reflect the potential consequences of healthcare-related harm: longer than expected length of stay (LOS) and unplanned readmission. Such indicators might be used to identify those patients in whom it is likely harm has occurred even if it had not been recorded. In this way the detection or yield of harm could be enhanced. Support for such an approach comes from RCRR studies that have shown adverse events are associated with longer LOS,^{3–5} though the direction of causality is unclear as a longer stay increases the risk of an error in care and subsequent harm.^{6–7} Similarly, a high rate of unplanned readmissions has been shown to be associated with harm having occurred.^{8–10}

Our aim was to identify ways of using routine hospital data to improve the efficiency of retrospective reviews of case records for identifying those patients who suffered avoidable severe harm. In this paper we focus on exploring the potential use of two indirect measures (long lengths of stay and early unplanned readmissions) in patients with one of four tracer conditions (acute myocardial infarction (AMI), bowel cancer surgery (elective and emergency) and hip replacement). These were selected to represent elective, urgent and emergency admissions for medical and surgical reasons.

With lengths of stay we have evaluated different thresholds for defining a long stay; while with early readmissions we have assessed a range of criteria. To evaluate different thresholds we assessed how they are associated with the presence of direct indicators of harm coded within the electronic care records. We then used this to understand the resource implications for choosing different screening criteria. Our specific analyses, therefore, focused on:

- ▶ The relationships between direct indicators of harm and:
 - Long lengths of stay.
 - The time between a patient is discharged from hospital and readmitted as an emergency.
 - The length of a readmission spell.
- ▶ How the presence of a direct indicator of harm affects the chances of a subsequent readmission.
- ▶ Primary diagnoses on readmission that reflect potential harm.

METHOD

Data

For this analysis we used hospital administrative data as reported in Hospital Episode Statistics (HES) for the 2014–2015 financial year. Diagnosis and procedure codes are based on the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) and the Office of Population Censuses and Surveys (OPCS) Classification of Interventions and Procedures, version 4 (OPCS-4). There are 20 diagnostic fields per episode: a primary diagnosis and up to 19 secondary diagnoses.

Categorising direct indicators of harm

A set of ICD-10 codes were selected as direct indicators of harm. These included complications and adverse reactions (T80-88; Y40-84) plus others identified from the literature,^{11–13} or through consultation with clinicians and clinical coders and other sources of guidance.^{14–15} Direct indicators of harm were divided into eight groups: complications and adverse reactions; thromboembolism; pneumonia; pressure sores; urinary tract infections; falls; fractures; post-procedural complications. Further details on these definitions are shown in online supplementary appendix 1.

Selection of patient cohorts

Given that the approach to using direct or indirect indicators might vary by the type of admission and condition, we selected four examples that represented both medical and surgical conditions and the urgency of the admission: (1) emergency: AMI; (2) urgent: non-elective bowel cancer surgery; (3) semiurgent: elective bowel cancer surgery; and (4) elective: hip replacement. A full list of ICD-10 and OPCS-4 codes for defining these cohorts is supplied in online supplementary appendix 2. Cohorts were restricted to adults (over 17 years), their first admissions in 2014–2015 to an acute NHS hospital in England for the relevant condition or surgery, and discharged alive. Admissions excluded day cases or regular day or night attenders.

Evaluating thresholds for unexpected long LOS

LOS was measured as the time between admission and discharge, ignoring transfers to other hospitals. The expected LOS of each patient was estimated using a linear regression model that controlled for age, sex, comorbidities, deprivation and emergency admissions in the previous 12 months. Comorbidities were measured using the Charlson Score and deprivation by quintiles of the Index of Multiple Deprivation (IMD). The expected LOS was estimated for each of the patient cohorts in 2013–2014 and the parameter estimates applied to the 2014–2015 population.

Thresholds for long lengths of stay were defined as multiples of the expected values, specifically two, three, four and five times. For each threshold, we investigated the association with the direct indicators of harm using a

linear regression model adjusting for age, sex, Charlson Score, IMD and number of emergency admissions in the previous year. We then evaluated the impact of different thresholds on the number of patient records that each trust would have to review in order to find instances of harm (positive predictive value; PPV) and the proportion of patients with harms reported in the spell that would be selected (sensitivity).

Unplanned readmission

Patients in each cohort were identified as having an unplanned readmission if their subsequent admission was an emergency and occurred either in 2014–2015 or 2015–2016. Screening criteria for readmissions were derived from combinations of:

- ▶ Time to readmission.
- ▶ Harm reported in the first episode of the readmission spell.
- ▶ Harm reported in the initial admission (present either in second or subsequent episodes).
- ▶ Primary diagnosis on readmission.
- ▶ Length of the readmission spell.

The relevance of the primary diagnosis on readmission to harm having occurred in the previous admission was determined by expert clinical review. The reviewers judged whether the primary diagnosis codes used in the dataset could represent healthcare-related harm. This was done for each cohort to allow for differences types of harm between the four cohorts (eg, conditions that are relevant for the hip replacement cohort may not be relevant for the AMI cohort).

As with lengths of stay, different options were evaluated in terms of proportions of case notes to be reviewed, and the sensitivity and positive predicted value associated with the occurrence of direct indicators of harm.

Patient involvement

There were two patients on the steering group for this study. One was recruited through a local hospital patient reference group and the other was a patient advisor for a charity auditing the care of acutely ill patients, who was recruited through contact with the charity. Both had experience of family illness and in the case of one representative, a family member who had experienced a healthcare-related harm. The two patient representatives contributed, through discussion at meetings, to the design of the study and suggested a range of possible harms that the study could look at. They also provided helpful input as to how study results might be effectively communicated to wider audiences.

RESULTS

Mean age was similar across the patient cohorts (68–70 years) and each cohort had similar distribution of socio-economic status (table 1). There were differences in sex: men made up 66% of the AMI cohort but only 40% of the elective hip replacements. There were also differences

in comorbidity: 45% of emergency bowel cancer patients had Charlson scores of four or more compared with 5% of patients receiving elective hip replacement.

What is the relationship between long lengths of stay and direct indicators of harm?

The median LOS differed between cohorts (table 2) and the distribution was highly positively skewed. The prevalence of harm increased with LOS (figure 1). For example, 94% of emergency bowel surgery patients staying longer than 50 days had experienced harm compared with 16% in those who stayed 5–9 days. Linear regression analysis found that nearly all categories of harm were significantly positively associated with LOS: some exceptions being fractures in emergency bowel cancer patients and hospital-acquired infections in all bowel cancer patients (online supplementary appendix 3, table 3.1).

What are the resource implications for choosing different screening criteria derived from LOS?

The impacts of different LOS thresholds on the numbers of patient notes that would be selected and on the sensitivity of detecting harm are shown in table 3.

Of all the patients with a direct indicator of harm, the proportion included in these subgroups (the sensitivity) decreases as the threshold rises, from 56% to 15%. At the same time, the PPV (the number of cases identified in each threshold that would actually have a harm code) increases. For example, for hip replacement the value rises from 10% among all patients to 51% for those staying three times longer than expected.

What is the relationship between the presence of a direct indicator of harm and time to emergency readmission?

Rates of readmission within 7 days for AMI and bowel surgery (6%–7%) are notably higher than for hip replacement (2.6%) (table 2). Approximately half of readmissions within 28 days occur within the first 7 days. More than half the surgical patients readmitted within 7 days had a direct indicator of harm compared with 21% in the AMI cohort (figure 2). With bowel surgery cases, these proportions decline as time to readmission increases but for hip replacement patients, the decline is only after 28 days and for AMI patients there is no association with time to readmission. This pattern among the surgical cohorts is specifically due to declines in ‘complications and adverse reactions’ which constitutes approximately 65% of harm across these groups in contrast to the AMI cohort where only 25% are ‘complications and adverse reactions’.

How does the presence of a direct indicator of harm affect the chances of a subsequent emergency readmission?

For individuals who have a direct indicator of harm reported in the initial spell, the 7 day readmission rates are higher than the overall 7 day rates for each cohort except those undergoing urgent bowel cancer surgery (table 4). After adjusting for age and likelihood of readmission (using Patients at Risk of Re-admission within 30

Table 1 Patient cohort characteristics

	Acute myocardial infarction		Urgent bowel surgery		Elective bowel surgery		Hip replacement	
	N	%	N	%	N	%	N	%
Sex								
Male	39887	66	1637	52	9856	58	21206	40
Female	20299	34	1490	48	7245	42	31532	60
Age group (years)								
<54	11017	18	471	15	1849	11	6651	13
55–64	12504	21	533	17	3567	21	10622	20
65–74	13935	23	801	26	5757	34	18345	35
75–84	14029	23	951	30	4839	28	14437	27
>85	8701	14	371	12	1089	6	2683	5
Charlson score								
0	24299	40	1037	33	8286	48	34203	65
1	9448	16	203	6	1739	10	4776	9
2	10597	18	307	10	2174	13	8135	15
3	6483	11	166	5	951	6	3238	6
4+	9359	16	1414	45	3951	23	2386	5
IMD quintiles								
1 (least deprived)	13148	22	551	18	2489	15	7203	14
2	12273	21	591	19	2975	18	9222	18
3	12129	20	678	22	3699	22	11380	22
4	11621	20	684	22	3959	23	12326	24
5 (most deprived)	10258	17	588	19	3815	23	11811	23
Admission in previous year	13470	22	980	31	4020	24	5221	10
Direct indicator of harm present during the hospital spell	8348	14	1080	35	4479	26	5317	10
Total	60186		3127		17101		52738	

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved. IMD, Index of Multiple Deprivation.

days Score¹⁶), the time to readmission was only related to the record of a direct indicator of harm in the initial spell for the AMI cohort (excluding cases where there was a

direct indicator of harm reported in both the initial spell and the readmission).

Table 2 Median and average lengths of stay and readmissions by condition

	Total number of patients in cohort	Lengths of stay				Readmissions within 7 days		Readmissions within 28 days	
		Median (days)	Mean (days)	Min (days)	Max (days)	Number of readmissions	Rate	Number of readmissions	Rate
Acute myocardial infarction	60186	4	6.5	0	412	4072	6.8%	8149	13.5%
Emergency bowel cancer surgery	3127	13	17.4	0	207	202	6.5%	445	14.2%
Elective bowel cancer surgery	17101	7	9.7	0	235	1086	6.4%	2057	12.0%
Elective hip replacement surgery	52738	4	4.8	0	174	1344	2.6%	2776	5.3%

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved.

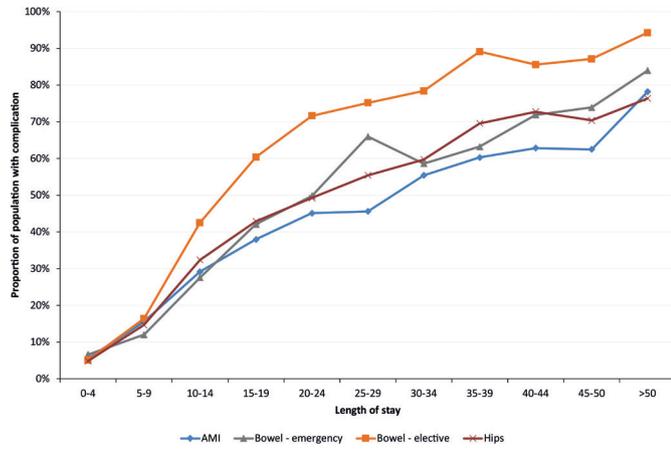


Figure 1 Prevalence of harm by length of stay. AMI, acute myocardial infarction.

How is the presence of a direct indicator of harm related to the length of a readmission spell?

The proportion being readmitted for more than 3 days varied by cohort: bowel surgery over 50%, AMI 44%, hip replacement 32%. The latter group have more patients who stay for less than a day (32% compared with 13% to 17% for the other cohorts). 44% of these readmissions are for conditions reported as 'other soft tissue disorders' and they also include all patients with a primary diagnosis of 'phlebitis and thrombophlebitis' (including deep vein thrombosis). The latter represent cases where patients are discharged quickly after the readmission to manage the condition in the community. Direct indicators of harm are significantly more prevalent when the

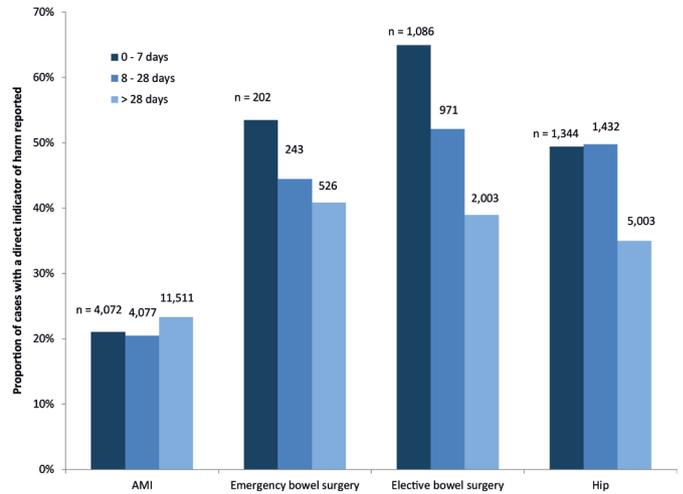


Figure 2 Proportions of cases with direct indicators of harm reported in the first episode of the readmission spell by time to readmission. AMI, acute myocardial infarction.

readmission LOS is longer than 3 days (table 5) ($p < 0.001$ for each cohort).

In how many readmissions is a potential harm suggested by the primary diagnosis?

Just over half of readmissions within 7 days for the AMI cohort were admitted with a primary diagnosis that was judged by expert review to be potentially related to harm (table 6). This compares with much higher proportions among the other cohorts, with nearly 99% of the hip surgery cohort having a diagnosis that could be potentially related to harm among that group (more details in

Table 3 Impact of different 'longer than expected LOS' thresholds on (1) the proportion of patients selected and (2) the sensitivity and positive predictive value (PPV) associated with detecting a direct indicator of harm

Threshold	All cases	Longer than expected	Two x longer than expected	Three x longer than expected
Acute myocardial infarction				
% selected (n)	100 (60 186)	28 (17 072)	9 (5664)	4 (2618)
Sensitivity	100%	56%	27%	15%
PPV	14%	27%	40%	47%
Urgent bowel surgery				
% selected (n)	100 (3127)	58 (1806)	20 (631)	9 (270)
Sensitivity	100%	81%	39%	19%
PPV	35%	48%	66%	74%
Elective bowel surgery				
% selected (n)	100 (17 101)	26 (4420)	7 (1163)	3 (468)
Sensitivity	100%	59%	21%	9%
PPV	26%	60%	82%	88%
Hip replacement				
% selected (n)	100 (52 738)	31 (16 528)	5 (2516)	2 (835)
Sensitivity	100%	60%	20%	8%
PPV	10%	19%	41%	51%

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved.

Table 4 Seven-day readmission proportions for patients with a direct indicator of harm in the second or subsequent episode of the initial spell

	A direct indicator of harm in the second or subsequent episode of the initial spell				
	Present			Overall 7 day readmission rate	P value associated with the presence of direct harm
	Number of patients	Number readmitted	Proportion readmitted within 7 days		
Acute myocardial infarction	2569	218	8.5%	6.8%	<0.001
Urgent bowel surgery	393	24	6.1%	6.5%	0.76
Elective bowel surgery	528	42	8.0%	6.4%	0.12
Hip replacement	150	8	5.3%	2.6%	0.06

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved.

online supplementary appendix 3, table 3.2). There were no significant differences in the proportions within 7 days from 8 to 28 days among the surgical cohorts. However, there are significant reductions in proportions among elective bowel surgery readmissions that occur after 28 days ($p<0.001$). Among the AMI cohort, the proportion among the earlier readmissions (50.9%) is significantly higher than among the later readmissions ($p<0.001$).

What are the resource implications for choosing different screening criteria derived from emergency readmissions?

Choices of criteria against which to select case records for review will depend on a trade-off between numbers of cases selected and proportion of harm that is found. [table 7](#) shows the outcomes of different criteria using 28 day readmissions as a baseline against which to compare proportions of notes selected and sensitivities. With the hip surgery and elective bowel replacement cohorts, given the majority of the primary diagnoses on readmission are associated with harm having occurred ([table 7](#)), restricting selection to these primary diagnoses (scenario C) makes little difference. Further limiting selection to cases where readmission lengths of stay exceed 3 days will

reduce the number of case records for review by 50% or more but will correspond to larger reductions in sensitivity (comparing scenario E with scenario C). Including any cases where direct indicators of harm are present, regardless of LOS and primary diagnosis will increase the PPV at the expense of having a larger proportion of notes to review.

DISCUSSION

Main findings

It is possible to derive criteria from hospital administrative data to select case records in order to find cases of severe hospital-related harm. Our findings suggest that adopting screening rules based on two indirect indicators (long lengths of stay and early readmission) has the potential to improve the targeting of case record reviews. The precise scale of any improvements is unclear until selection criteria have been tested against the outcomes of such reviews.

The selection of LOS thresholds for screening could have a significant impact on the yield of cases of harm. For

Table 5 Proportion of direct indicators of harm by length of readmission spell for readmissions within 7 days

Length of stay (days)	Acute myocardial infarction		Bowel surgery		Hip replacement	
	Number of readmissions	Proportion with direct indicator of harm (95% CI)	Number of readmissions	Proportion with direct indicator of harm (95% CI)	Number of readmissions	Proportion with direct indicator of harm (95% CI)
0	560	14.1% (11.2% to 17.0%)	211	62.6% (56.1% to 69.1%)	427	34.7% (30.2% to 39.2%)
1	852	12.4% (10.2% to 14.6%)	158	50.0% (42.2% to 57.8%)	248	49.2% (43.0% to 55.4%)
2	492	13.8% (10.8% to 16.9%)	123	61.0% (52.4% to 69.6%)	127	48.8% (40.1% to 57.5%)
3	376	19.9% (15.9% to 23.9%)	127	53.5% (44.8% to 62.2%)	112	50.0% (40.7% to 59.3%)
>3	1792	29.6% (27.5% to 31.7%)	669	68.6% (65.1% to 72.1%)	430	64.2% (60.0% to 68.8%)

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved.

Table 6 Proportions of patients readmitted with a primary diagnosis that is potentially related to harm

Cohort	Numbers with potential harm-related primary diagnosis					
	Readmissions within 7 days		Readmissions between 8 and 28 days		Readmissions after 28 days	
	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)
AMI	2071	50.9 (49.4 to 52.4)	1802	44.2 (42.7 to 45.7)	5024	43.6 (42.7 to 44.5)
Urgent bowel surgery	169	83.7 (78.6 to 88.8)	204	84.0 (79.3 to 88.6)	432	82.1 (78.9 to 85.4)
Elective bowel surgery	1019	93.8 (92.4 to 95.3)	915	94.2 (92.8 to 95.7)	1807	90.2 (88.9 to 91.5)
Hip replacement	1325	98.6 (98.0 to 99.2)	1417	99.0 (98.5 to 99.5)	5195	98.0 (97.6 to 98.4)

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved.

AMI, acute myocardial infarction.

example, over half those who stayed at least three times longer than expected had a direct indicator of harm. The PPV of the screen increases across the thresholds, such that the number of cases identified as having a direct indicator of harm as a proportion of all cases examined

increases. By manipulating LOS threshold, choices can be made in relation to the trade-off between the number of cases that will actually have a harm code present at that threshold and the proportion of all the harm that will be found if only those cases are investigated.

Table 7 Implications of different scenarios for selecting case notes for readmitted patients in terms of the proportion of 28 day readmissions selected and the sensitivity and positive predictive value (PPV) associated with detecting a direct indicator of harm occurring within 28 days

Scenario	A All patients readmitted within 28 days	B All patients readmitted within 7 days	C 7 day readmissions: only primary diagnoses associated with potential harm	D 7 day readmissions: primary diagnoses associated with potential harm or other reported direct indicators of harm	E 7 day readmissions: primary diagnoses associated with potential harm if length of readmission spell > three days	F 7 day readmissions: primary diagnoses associated with potential harm if length of readmission spell > three days, or other reported direct indicators of harm regardless of length of stay
Acute myocardial infarction						
% selected (n)*	100 (8149)	50 (4072)	25 (2071)	28 (2313)	13 (1084)	19 (1544)
Sensitivity†	100%	51%	36%	51%	24%	51%
PPV‡	21%	21%	30%	37%	37%	56%
Urgent bowel surgery						
% selected (n)*	100 (445)	45 (202)	38 (169)	41 (181)	18 (81)	32 (141)
Sensitivity	100%	50%	44%	50%	22%	50%
PPV	49%	53%	57%	60%	59%	77%
Elective bowel surgery						
% selected (n)*	100 (2057)	53 (1086)	50 (1019)	50 (1038)	26 (533)	41 (849)
Sensitivity	100%	58%	57%	58%	32%	58%
PPV	59%	65%	67%	68%	73%	83%
Hip replacement						
% selected (n)*	100 (2776)	48 (1344)	48 (1325)	48 (1330)	15 (420)	29 (812)
Sensitivity	100%	48%	48%	48%	20%	48%
PPV	50%	49%	50%	50%	65%	82%

*Proportions are of all readmissions within 28 days.

†Proportion of direct indicators of harm occurring in readmissions within 28 days.

‡Proportion of selected cases that have a direct indicator of harm.

Hospital Episode Statistics data ©2017, reused with the permission of the Health & Social Care Information Centre. All rights reserved.

Selection of the time to readmission has an effect on the yield of potential cases of harm but it varies by condition. At least 50% or more of the surgical patients had a direct indicator of harm if readmitted within 7 days, compared with 21% in the AMI cohort. With bowel surgery cases, these proportions decline as time to readmission increases but for hip replacement patients, the decline is only after 28 days and for patients with AMI there is no association with time to readmission. This suggests that the sampling window for the latter two conditions could be extended to 28 days without significant impact, with the added benefit of increasing the number of patients in the hip replacement cohort where there are relatively few readmissions. The lack of relationships between a harm code found in the initial admission with time to readmission suggests that the occurrence of harms in the initial episode may not be useful as a criterion for selecting case records, except, perhaps, for patients with AMI. However, because we were not able to identify individuals who had died outside hospital soon after the initial spell this analysis may underestimate subsequent outcomes after discharge following a harm.

For primary diagnoses on readmission deemed to be potentially associated with harm, there were higher frequencies among the surgical cohorts with between 80% and 100% of readmission primary diagnoses identified by clinical reviewers being potentially associated with harm. For the AMI cohort, the corresponding proportion was around 50%. This suggests that the nature of the primary readmission diagnosis can be useful as a further criterion for selecting case records and this approach would have the greatest impact on the AMI cohort.

Our assessments of thresholds used PPVs and sensitivity as we were interested in the value of case note review in revealing a harm and an indication of how effective they are at detecting all harms that may have occurred. We could also have used specificity and negative predictive values, but considered them less useful in this context.

Other literature

Previous RCRR studies estimated that the proportion of inpatients with an adverse event ranged from 3.8% to 16.6%.¹⁷ Across the four cohorts, we found higher proportions of harm codes. However, the conditions we studied were chosen to highlight different admission types and were not representative of all conditions. Similar rates of harm in bowel cancer patients (between 20% and 40%) have been found in previous studies.¹⁸ A recent Dutch study found a higher proportion of harm in patients admitted with AMI, between 13.3% and 29.9%.¹⁹ The harm in this study was found using an audit tool to screen electronic patient records which could account for a greater proportion of harm being uncovered. National clinical audits suggest that the rate of complications after percutaneous coronary interventions is around 9% in England,²⁰ and after total hip replacement are about 1% for infection and venous thromboembolism and 3%–4%

for dislocation.²¹ This is consistent with the lower incidence of harm that we found in this group.

Our study is the first to look at the relationships between different LOS thresholds and a variety of readmission characteristics and coded harm in hospital administrative records in the UK. The Dutch have used a threshold based indicator, unexpectedly long LOS (UL-LOS), defined as the percentage of clinically admitted patients with an actual hospital stay that is more than 50% longer than expected, as a generic indicator of hospital safety for a number of years.²² Cihangir *et al*²³ found a significant positive correlation between UL-LOS and another indicator of potentially poor quality care, the hospital standardised mortality ratio ($r=0.44$, $p<0.001$) in hospital administrative data from two-thirds of Dutch hospitals. In a small, single site validation study the authors found that in 85 out of 191 patients with colorectal cancer with UL-LOS, 43 (51%) had one or more adverse events, compared with 9% (4 out of 44) in the non-unexpected long LOS group.²⁴

Limitations

There are a number of limitations to be considered when interpreting our findings. First, our estimates of harm, based on our inclusive approach, are inflated by an element of double counting when the same harms are coded under more than one of the harm categories. This inflation is compounded by the fact that we inevitably include a number of conditions that were present on admission without the routine application of 'present on admission' codes. Harm codes such as pneumonia or urinary tract infection, which occur more commonly in emergency medical patients, are also more difficult to attribute to healthcare-related processes. It is also known that hospitals vary in the way they use complication and adverse reaction codes which has the potential to introduce further bias in measurement.¹

Second, one of the main limitations of developing screening approaches is the accuracy and completeness of the routine data.¹²⁵ This limitation is particularly important to consider in this study which used harm codes for the internal validation of indirect indicators of harm derived from the same routine hospital administrative data source. As our approach to harm code definitions was inclusive in an attempt to increase sensitivity, it is likely that our estimates are inflated which may have biased assessment of the performance of indirect indicators as screening tools.

Third, to assess the feasibility of using indirect indicators to detect cases of harm, we have relied on direct indicators. However, we cannot know the relationship with other types of harm that are not so easily identifiable within routine data, and whether approaches that would work for the detection of harm codes we identified would work more generally. Our analytic approach can indicate that a patient may have experienced harm and the patterns of harm among groups of patients with differing conditions and across an organisation but it cannot

confirm if that harm was healthcare-related, its severity or its avoidability without recourse to case record review.

Finally, not all long lengths of stay reflect a patient's acute care needs as there may be several days when a patient is awaiting discharge. However, it has not been possible to distinguish these from the data.

Implications

Screening for harm using routine data allows large numbers of records to be rapidly processed with minimal resources required. The Global Trigger Tool has also been developed as a harm screening tool.²⁶ However using this tool to identify triggers linked to harm, case records need to be individually screened, which is usually done manually creating a more resource intense process. Our approach to screening using long lengths of stay or early readmissions would identify a substantial number of patients for case record review if extended across all patients. In 2014–2015, there were approximately 16 million admissions to all NHS acute trusts in England,²⁷ yet the cohorts we included in this analysis comprised less than 1%. If a threshold of twice the expected LOS was used for screening, we estimate this would have resulted in 9974 case record reviews in 2014–2015 across the four cohorts, which equates to about 70 per trust. If we assume a similar LOS distribution across all hospital admissions, this scales up to around 7000 reviews per year in an average sized trust. Furthermore, having increased the sensitivity of the screening process at the expense of specificity, many of these cases would contain no evidence of healthcare-related harm.

In practice, therefore, this suggests that one approach might be for Trusts to decide how many reviews they are going to do in advance and then select random samples of cases that satisfy the screening criteria. Consideration of how to adapt or create algorithms applicable to wider patient populations would also be required.

Contributors HH, NB, CSJ and JVM designed the study. CSJ, NCO and KC carried out the analyses. NB, CSJ and HH wrote the first draft of the manuscript. All authors provided input and approved the final version for submission.

Funding This work was supported by Department of Health Policy Research Programme (PR-R9-0114-14001).

Disclaimer The views expressed in this publication are those of the authors and not necessarily those of the Department of Health. The funders of the study had no role in the study design; data collection, analysis, and interpretation; or composition of the report.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by North West- Lancaster Research Ethics Committee (15/NW/0941). As data were pseudoanonymised individual patient consent was not required. Hospital Episode Statistics data ©2017, re-used with the permission of the Health & Social Care Information Centre. All rights reserved.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Further details on statistical models and definitions are available from the Nuffield Trust at research@nuffieldtrust.org.uk.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Aylin P, Tanna S, Bottle A, *et al*. How often are adverse events reported in English hospital statistics? *BMJ* 2004;329:369.
2. Australian Institute for Health and Welfare. *Reporting of adverse event in routinely collected data sets in Australia*: Australian Institute for Health and Welfare, 2001.
3. Sari AB, Sheldon TA, Cracknell A, *et al*. Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital. *Qual Saf Health Care* 2007;16:434–9.
4. Baker GR, Norton PG, Flintoft V, *et al*. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004;170:1678–86.
5. Davis P, Lay-Yee R, Briant R, *et al*. Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J* 2002;115:U271.
6. Asher RA. The dangers of going to bed. *Crit Care Update* 1983;10:40.
7. Kandula PV, Wenzel RP. Postoperative wound infection after total abdominal hysterectomy: a controlled study of the increased duration of hospital stay and trends in postoperative wound infection. *Am J Infect Control* 1993;21:201–4.
8. Care Quality Commission Intelligent monitoring. *NHS acute hospitals- indicators and methodology guidance*: Care Quality Commission, 2015.
9. Nolte E, McKee M, Wait S. *Chapter 2. Describing and evaluating health systems*. Maidenhead: Open University Press, 2005.
10. Leng GC, Walsh D, Fowkes FG, *et al*. Is the emergency readmission rate a valid outcome indicator? *Qual Health Care* 1999;8:234–8.
11. Ghali WA, Pincus HA, Southern DA, *et al*. ICD-11 for quality and safety: overview of the WHO Quality and Safety Topic Advisory Group. *Int J Qual Health Care* 2013;25:621–5.
12. Brand CA, Sundararajan V. A 10-year cohort study of the burden and risk of in-hospital falls and fractures using routinely collected hospital data. *Qual Saf Health Care* 2010;19:e51.
13. Blunt I, Bardsley M, Grove A, *et al*. Classifying emergency 30-day readmissions in England using routine hospital data 2004–2010: what is the scope for reduction? *Emerg Med J* 2015;32:44–50.
14. Health and Social Care Information Centre. *National Clinical Coding Standards ICD-10*. 4th edn. Addendum: Accurate data for quality information. Health and Social Care Information Centre, 2017.
15. NHS Digital. *HES Data Dictionary: Admitted Patient Care*: NHS Digital, 2016.
16. 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;2:e001667.
17. de Vries EN, Ramrattan MA, Smorenburg SM, *et al*. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216–23.
18. Alves A, Panis Y, Mathieu P, *et al*. Postoperative mortality and morbidity in French patients undergoing colorectal surgery: results of a prospective multicenter study. *Arch Surg* 2005;140:278–83.
19. Eindhoven DC, Borleffs CJ, Dietz MF, *et al*. Design and reliability of a specific instrument to evaluate patient safety for patients with acute myocardial infarction treated in a predefined care track: a retrospective patient record review study in a single tertiary hospital in the Netherlands. *BMJ Open* 2017;7:e014360.
20. Banning AP, Baumbach A, Blackman D, *et al*. Percutaneous coronary intervention in the UK: recommendations for good practice 2015. *Heart* 2015;101(Suppl 3):1–13.
21. National Joint Registry Editorial Board. *National Joint Registry for England, Wales, Northern Ireland and Isle of Man 13th Annual Report*: National Joint Registry, 2016.
22. Ghielen J, Cihangir S, Hekkert K, *et al*. Can differences in length of stay between Dutch university hospitals and other hospitals be explained by patient characteristics? A cross-sectional study. *BMJ Open* 2019;9:e021851.
23. Borghans I, Hekkert KD, den Ouden L, *et al*. Unexpectedly long hospital stays as an indicator of risk of unsafe care: an exploratory study. *BMJ Open* 2014;4:e004773.
24. Cihangir S, Borghans I, Hekkert K, *et al*. A pilot study on record reviewing with a priori patient selection. *BMJ Open* 2013;3:e003034.
25. Raftery J, Roderick P, Stevens A. Potential use of routine databases in health technology assessment. *Health Technol Assess* 2005;9:1–92.



26. Griffin F, Resar R. *IHI Global Trigger Tool for measuring adverse events*. 2nd edn. Cambridge, Massachusetts: Institute for Healthcare Improvement, 2009.
27. Health and Social Care Information Centre. Hospital Episode Statistics: Admitted Patient Care, England 2014-15. 2015 <https://files.digital.nhs.uk/publicationimport/pub19xxx/pub19124/hosp-epis-stat-admi-summ-rep-2014-15-rep.pdf>.