



OPEN ACCESS



Open Access
Scan to access more
free content

The Patient-Reported Incident in Hospital Instrument (PRIH-I): assessments of data quality, test–retest reliability and hospital-level reliability

Oyvind Bjertnaes, Kjersti Eeg Skudal, Hilde Hestad Iversen, Anne Karin Lindahl

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2012-001756>).

Department for Quality and Patient Safety, The Norwegian Knowledge Centre for the Health Services, Oslo, Norway

Correspondence to

Dr Oyvind Bjertnaes, Department for Quality and Patient Safety, The Norwegian Knowledge Centre for the Health Services, Boks 7004 St Olavs plass, Oslo 0130, Norway; oan@kunnskapssenteret.no

Received 13 December 2012

Revised 4 April 2013

Accepted 11 April 2013

Published Online First

14 May 2013

ABSTRACT

Background The objective of this study was to test the data quality, test–retest reliability and hospital-level reliability of the Patient-Reported Incident in Hospital Instrument (PRIH-I).

Methods 13 incident questions were included in a national patient-experience survey in Norway during the spring of 2011. All questions and a composite incident index were assessed by calculating missing-item rates, test–retest reliability and hospital-level reliability.

A multivariate linear regression on a global item regarding incorrect treatment was used to assess the main sources of variation in patient-perceived incorrect treatment at hospitals.

Results Five of the 13 patient-incident questions had a missing-item rate of >20%. Only one item met the criterion of 0.7 for test–retest reliability (wrong or delayed diagnosis), seven items had a score of >0.5, while the remainder had a reliability score of <0.5. However, the reliability was >0.7 for six of 10 items tested at the hospital level, and >0.6 for the remaining four items. A patient-incident index based on 12 of the incident items had no missing data, the test–retest reliability was 0.6 and the hospital-level reliability was 0.85.

Conclusions The PRIH-I comprises 13 questions about patient-perceived incidents in hospitals, and can be easily and cost-effectively included in national patient-experience surveys with an acceptable increase in respondent burden. Although the missing-item rate and test–retest reliability were poor for several items, the hospital-level reliability was satisfactory for most of the items. The incident items contribute to a patient-reported incident index, with excellent data quality and hospital-level reliability.

BACKGROUND

Patient experiences and satisfaction are important determinants of the quality of healthcare.¹ Patient experiences constitute a core dimension of quality in the Organization for Economic Co-operation and Development quality indicator project, together with effectiveness and safety.² Patient experiences and satisfaction are often measured with the aid of surveys^{3–4} that have traditionally focused on patients' non-medical needs and preferences, such as the need for adequate information, communication and organisation.⁵ The safety and effectiveness components of quality have been measured using other approaches, such as mortality assessments,⁶ the Global Trigger Tool⁷ and patient-reported outcome measures.⁸ However, there is a growing interest in involving patients in the evaluation of their safety.^{9–11} Important reasons for this are an increasing focus on patient participation and empowerment in Western healthcare systems, and a broader and more complex view of safety dimensions and perspectives.

Three reviews of the literature have recently been published,^{9–11} two of which have a particular focus on patient reporting of safety incidents in the hospital setting.^{9–10} According to these reviews, there is currently a positive and optimistic view in the literature regarding safety reporting by patients, but also concerns about methodology and the lack of connection between patient-safety reporting systems and quality improvement.

To cite: Bjertnaes O, Skudal KE, Iversen HH, *et al.* *BMJ Qual Saf* 2013;**22**:743–751.

One of the reviews identified 17 relevant publications, and found that the healthcare setting, method of reporting, time span, terminology, criteria for assessment and response rate varied considerably between them.¹¹ Another review also found considerable variation in focus, design and analysis between the 13 included studies, and concluded that from an academic perspective, patient reporting is in its infancy.⁹ The previous studies have not adequately documented the psychometric properties of the measurement instruments nor tested the reliability of patient-safety questions as quality indicators. Since patients struggle to understand patient-safety terminology,¹¹ it seems especially important to test the survey questions both cognitively and using psychometric methods. Furthermore, the need to move patients' safety reporting from the research domain to clinical governance⁹ includes an assessment of the usefulness and scientific adequacy of using such reports as quality indicators.

A national patient-safety campaign started in Norway on 1 January 2011. One aim for the campaign was to measure patient safety over time at different levels, including measurements of adverse events, patient-safety culture and patient reports of patient safety. Questions regarding patient-perceived incidents in hospitals were developed and tested, and these questions were included in the national patient-experience survey in 2011. The response rate for this survey has declined over time, and so the goal was to include a short instrument with little additional response burden. In addition, the new patient-incident questions to be used in the patient-safety campaign were to be included in future national reports on patient experiences in hospitals. Consequently, the incident questions were subjected to similar psychometric testing as the other experience questions. In Norway, patient-experience scales are aggregated to the hospital level and published as quality indicators in the national quality indicator system. Therefore, the new incident items also had to be tested as patient-perceived incident indicators at the hospital level.

Our standard development and validation process was the starting point for this project,^{12–15} including multiple activities to ensure content validity in the development phase, and a set of psychometric tests in the quantitative phase to assess instrument quality. Adjustments had to be made because of time restrictions and the measurement construct, the latter in particular resulting in substantial changes. Patient-perceived incidents are usually measured with single items, as opposed to the traditional multi-item scale approach in psychometrics. We chose to use the single-item approach because safety incidents are concrete factors, which are not well suited for the latent construct approach in multi-item scaling. However, this meant that core elements of the psychometric evaluation methods, such as factor analysis and internal consistency reliability, were inappropriate.

A revised evaluation approach was thus constructed, with a focus on data quality, test–retest reliability and hospital-level reliability. Hence, the evaluation approach assessed both traditional measures of reliability and the scientific adequacy of using patient-incident questions as quality indicators. The conceptual starting point was patient-safety outcomes or incidents, as perceived by patients. The national patient-reported experience questionnaire already included one general item about incorrect treatment, and we wanted to include more detailed questions about types of incidents. In addition, broader safety dimensions like trust in doctors and information received were already part of the national patient-experience questionnaire. We aimed to include important safety issues in hospitals, but at the same time were sensitive to the fact that most patients lack medical knowledge and might find safety incident reporting difficult. The lack of a comprehensive theoretical model implied that conceptual development and refinement were expected as a result of the qualitative development activities.

The objective of this study was to document the development, psychometric testing and hospital-level reliability testing of the Patient-Reported Incident in Hospital Instrument (PRIH-I). Preliminary questions were tested in cognitive interviews with 19 patients, and ultimately 13 incident questions were included in a national patient-experience survey in Norway during the spring of 2011.

METHODS

Development and cognitive testing of incident questions

The patient-incident questions were developed based on a review of the literature, interviews with internal experts and health personnel, and meetings in an internal reference group for the project. Two systematic reviews of the literature were identified^{10 11} in addition to other relevant national and international publications. Our review identified important safety domains, showed the importance of having supplemental open-ended questions, and also the importance of using lay language and cognitive testing with patients. In addition, the conceptual restriction to outcomes or incidents was challenged, since the review indicated the importance of both prevention of safety incidents and health personnel's actions after an incident has occurred.

We conducted five semistructured interviews with internal experts and health personnel with a focus on relevant topics for question generation and considerations of which topics patients are able to answer. Many topics were the same as identified in the review, and advice was also given on topics to exclude from the questionnaire. For instance, this related to psychological and socio-economical problems following an incident (the cause could both be the underlying disease and the incident), and failures with technical

equipment since technical issues are difficult to evaluate for patients. Furthermore, broader safety aspects like confidence in doctors and information were stressed, but these issues were already included in the national patient-experience questionnaire. The interviews showed that prevention of incidents also should be included, in line with findings from the review.

The findings from the review and the qualitative interviews were presented for the internal reference group, which supported the expansion to prevention of incidents and questions on how health personnel handle mistakes after they have occurred. The reference group was asked to rate the three most important safety incident domains and the consensus in the group was that the following topics were the most important: (i) wrong medications or other medication-related errors; (ii) administrative errors; and (iii) hospital infections. A set of safety questions was presented for the reference group, following some revisions and a final agreement on a set of preliminary questions for further testing.

The preliminary set of questions was tested in cognitive interviews with 19 patients visiting a university hospital in Norway. The interviews demonstrated that many patients felt unable to judge certain safety aspects, which meant that some questions had to include a 'do not know' option. The interviews also demonstrated that some of the safety questions were irrelevant for many patients, such as medication-related errors for patients not using medicines. Many of the patients perceived that no special incidents occurred, implying that a range of questions had to be answered even though they had nothing particularly negative to report about patient safety. This meant that some questions had to include a 'not applicable' option.

Data collection in the national survey

The details of the national survey have been described elsewhere.¹⁶ In short, it included 400 randomly selected adult inpatients who were discharged from each hospital in Norway between 1 March and 22 May 2011. Of these, 744 patients were not eligible, so that a final cohort of 23 420 patients was included in the study. The inclusion period was divided into three 4-week groups, meaning that patients received the questionnaire approximately 1–5 weeks following discharge. We wanted experiences to be as fresh as possible, but the timing question had to take account of practical issues relating to the fact that 61 hospitals had to transfer data to the Knowledge centre for each 4-week group. The response rate was 46.4%. A retest questionnaire was also mailed to 270 consenting patients approximately 1 week after their first reply for the purpose of assessing test–retest reliability; 163 patients returned the retest questionnaire (60.4%).

The Data Inspectorate and the Norwegian Ministry of Health and Care Services approved the survey.

Questionnaire used in the national survey

Based on the development project, 13 patient-incident questions were included in the patient-experience questionnaire.^{17 18} The questionnaire used in the national survey comprised 73 closed-ended items. Most experience items had a 5-point response format ranging from 1 ('not at all') to 5 ('to a very large extent'). A total of 35 items related to patient experiences with structures, processes and outcomes of healthcare were aggregated to the 10 quality indicators in the national report:¹⁹ waiting time (one item), physical hospital standard (six items), next of kin (two items), organisation (four items), doctor services (seven items), nursing services (seven items), information (three items), discharge planning (two items), cooperation with other health services (two items) and incorrect treatment (one item). The questionnaire also included an open-ended question on the last page asking for comments about their hospital stay or the questionnaire itself, and probed for information regarding errors or unnecessary problems during or after their hospital stay.

The 13 patient-incident questions and response categories are presented in online supplementary appendix 1. The topics covered were safety incidents, type of mistake, safety communication, medicine lists, infections, safety actions by health personnel such as hand hygiene and control of identity, and satisfaction with how health personnel handled the mistake after it occurred. Seven of 13 patient-safety questions had the same 5-point response format as described above, three questions had a 3-point format ('no', 'once' and 'more than once') and three questions had two response categories ('yes' and 'no').

An incident index consisting of 12 safety items was computed. The item about bringing a medicine list at admission to hospital was excluded because it partly relates to processes outside hospital control. For items with 5-point scales, the three worst response categories were given a value of 1, and the others were assigned a value of 0. For items with 3-point scales, the two worst response categories were given a value of 1, and the remaining was assigned a value of 0. For items with two response categories, the negative and positive response categories were given values of 1 and 0, respectively. Therefore, the incident index varied from 0 to 12, with higher numbers indicating more patient-perceived incidents.

Statistical analysis

Descriptives, missing-item rates and test–retest reliability were calculated for all patient-incident items and the incident index. The intraclass correlation coefficient was used as an estimate of the test–retest reliability for continuous variables, and the κ statistic for dichotomous variables. A widely accepted criterion is that the estimated coefficients should exceed 0.7.²⁰

Hospital-level reliability concerns the proportion of the variance in hospital-level mean scores that is due to true differences between hospitals, as opposed to differences that might be due to sampling errors.²¹ Hospital-level reliability was estimated for items with 3- and 5-point response scales using one-way Analysis of variance (ANOVA) to distinguish the between-hospital variance from the within-hospital variance. The intraclass correlation coefficient is calculated as the variation between hospitals divided by the total variation. The estimated reliability at the hospital level in the present study can be used to estimate how large samples in other studies would need to achieve specified levels of hospital reliability, such as how many respondents are needed to reach the standard reliability criterion of 0.7. These estimates were calculated using the Spearman–Brown prophecy formula.

Multivariate linear regression analysis was used to assess sources of variation in a global item about incorrect treatment, and included all specific safety items as predictors. The qualitative development work identified the importance of other safety-related experiences for patients' feeling of safety in hospitals, and these were also included as predictors in the regression: information (scale, three items), doctor services (scale, seven items), nursing services (scale, seven items) and organisation (three items). To avoid extensive loss of cases and generalisability, items with a large amount of missing data were recoded and

included in the regression as dummy variables. For example, the unnecessary damage related to surgery was recoded into two dummy variables: (i) patients answering 'not at all' or 'to a small extent' and (ii) patients not responding to the item. The reference group was patients answering 'to some extent', 'to a large extent' or 'to a very large extent', and regression coefficients should consequently be interpreted as the partial effect of being in the respective group compared with the reference group.

All analyses were conducted using SPSS (V.15.0), except for reliability analysis at the hospital level, for which Microsoft Office Excel 2007 was used.

RESULTS

Of the cohort of hospital patients, 10.5% perceived that they had been incorrectly treated by the hospital (table 1) 'to some extent' (6.4%), 'to a large extent' (2.3%) or 'to a very large extent' (1.8%). Most items had a large majority of responses in the most positive response category, indicating a high level of patient-perceived safety in the hospital. The most negative responses were found for questions about control of medicine lists, satisfaction with how health personnel handled the mistake or problem when it occurred, and transfer of important information between health personnel. The mean score for the patient-incident index was 1.21 (SD 1.69).

Table 1 Valid n, frequencies and means for patient-incident questions and the incident index (n=10 514)

	Valid n	Response category (%)					Mean (SD)
		1	2	3	4	5	
Questions with five response categories (1, 'not at all' to 5, 'to a very large extent')							
1. Do you believe you were incorrectly treated by the hospital in any way (as far as you yourself are able to judge)?	10 139	79.4	10.0	6.4	2.3	1.8	1.37 (0.85)
2. Did you get a wrong or delayed diagnosis in connection with your hospital stay?	7890	83.9	7.2	5.6	1.8	1.5	1.30 (0.78)
3. Did you suffer any unnecessary injury or unnecessary problem as a result of a surgical procedure or examination?	7291	84.1	5.7	4.7	2.6	2.9	1.35 (0.91)
4. Did you receive the wrong medication, or were you incorrectly medicated in any other way?	8481	90.0	4.6	3.4	1.1	0.8	1.18 (0.62)
5. Did you experience insufficient hand hygiene (hand washing) among the staff?	8849	86.7	8.9	3.0	0.8	0.6	1.20 (0.59)
6. If you experienced mistakes or unnecessary problems in connection with your hospital stay, did the staff handle the mistake or problem in a satisfactory way?	3056	19.7	9.2	15.4	32.4	23.3	3.30 (1.43)
7. Did you experience that important information about you had reached the right people?	9863	3.0	5.6	18.6	50.9	21.8	3.83 (0.93)
Questions with three response categories (1, 'no'; 2, 'once'; 3, 'more than once')							
8. Did you experience any administrative mistakes during your hospital stay?	10 195	87.0	9.6	3.5	–	–	1.17 (0.46)
9. Did you experience that the staff forgot to check your identity?	10 229	96.8	2.0	1.2	–	–	1.04 (0.26)
10. Did you experience that the staff forgot to give you important information?	10 191	92.4	5.0	2.6	–	–	1.10 (0.38)
Questions with two response categories							
11. Did you get an infection in connection with your hospital stay? (1, 'no'; 2, 'yes')	9650	90.1	9.9	–	–	–	–
12. Was an updated list of your medications reviewed with you when you were discharged from hospital? (1, 'yes'; 2, 'no')	5926	67.9	32.1	–	–	–	–
13. When you were admitted, did you bring along an updated list of medications you were taking? (1, 'yes'; 2, 'no')	7779	72.7	27.3	–	–	–	–
<i>Patient-incident index (range 0–12, sum of questions 1–12)</i>	10 514	–	–	–	–	–	1.21 (1.69)

Table 2 Missing-item rate, test–retest reliability and hospital-level reliability for patient-incident questions and incident index (n=10 514)

	Missing item (%)	Test–retest reliability (n=163)	Hospital-level reliability				
			ICC	Reliability*	n For rel. 0.7†	n For rel. 0.8†	n For rel. 0.9†
Questions with five response categories (1, 'not at all' to 5, 'to a very large extent')							
1. Do you believe you were incorrectly treated by the hospital in any way (as far as you yourself are able to judge)?	3.6	0.64	0.015	0.724	153	263	591
2. Did you get a wrong or delayed diagnosis in connection with your hospital stay?	25.0	0.72	0.024	0.760	95	163	366
3. Did you suffer any unnecessary injury or unnecessary problem as a result of a surgical procedure or examination?	30.7	0.56	0.020	0.704	117	198	446
4. Did you receive the wrong medication, or were you incorrectly medicated in any other way?	19.4	0.39	0.014	0.660	167	286	643
5. Did you experience insufficient hand hygiene (hand washing) among the staff?	15.8	0.54	0.014	0.673	164	282	634
6. If you experienced mistakes or unnecessary problems in connection with your hospital stay, did the staff handle the mistake or problem in a satisfactory way?	70.9	0.36	0.032	0.623	71	121	272
7. Did you experience that important information about you had reached the right people?	6.2	0.38	0.034	0.851	66	114	256
Questions with three response categories (1, 'no'; 2, 'once'; 3, 'more than once')							
8. Did you experience any administrative mistakes during your hospital stay?	3.0	0.54	0.022	0.795	104	178	400
9. Did you experience that the staff forgot to check your identity?	2.7	0.51	0.009	0.604	263	451	1014
10. Did you experience that the staff forgot to give you important information?	3.1	0.16	0.014	0.709	164	282	634
Questions with two response categories							
11. Did you get an infection in connection with your hospital stay? (1, 'no'; 2, 'yes')	8.3	0.44	–	–	–	–	–
12. Was an updated list of your medications reviewed with you when you were discharged from hospital? (1, 'yes'; 2, 'no')	26.0	0.67	–	–	–	–	–
13. When you were admitted, did you bring along an updated list of medications you were taking? (1, 'yes'; 2, 'no')	43.6	0.61	–	–	–	–	–
<i>Patient-incident index (range 0–12, sum of questions 1–12)</i>	0.0	0.61	0.031	0.845	73	126	283

*Reliability refers to observed reliability given the sample size.

†n For rel. 0.7/0.8/0.9 indicates the sample size per hospital needed to obtain these three reliability levels.

ICC, intraclass correlation coefficient.

The missing-item rate exceeded 20% for five of the 13 patient-incident questions (table 2). The item about satisfaction with how health personnel handled the mistake or problem was unanswered by 70.9% of patients, while item-missing was 43.6% for the question about bringing an updated medicine list at admission. The missing-item rates were the lowest for questions regarding control of identity (2.7%), administrative mistakes (3.0%) and receiving important information (3.1%). There were no missing cases for the incident index.

Only one item met the criterion of 0.7 for test–retest reliability, namely, that about wrong or delayed diagnosis (table 2). Seven items had a test–retest reliability greater than 0.5, ranging from 0.67 for the item about updated medicine list check at discharge to

0.51 for the item about control of identity. The remaining items had a test–retest reliability rating of less than 0.5, ranging from 0.44 for hospital infections to 0.16 for the item about receiving important information from hospital staff. The test–retest reliability rating for the incident index was 0.61.

The reliability rating exceeded 0.7 for six out of 10 continuous items at the hospital level, ranging from 0.85 for the item about transfer of important information between health personnel to 0.70 for the item about unnecessary injury/problem related to surgery. The remaining four continuous items had a hospital-level reliability of between 0.6 and 0.7, while the hospital reliability was 0.85 for the incident index. With the exception of the item about control of identity, fewer than 200 respondents were needed for each

hospital to reach the reliability criterion of 0.7. A reliability criterion of 0.9 would require very large hospital samples for all incident questions (range 256–1014), but the incident index and two items reach 0.9 with less than 300 respondents. Figure 1 shows the mean score for the patient-incident index by hospital, which varied from 0.54 for the best hospital to 1.68 for the worst.

Multivariate linear regression showed that unnecessary injury following surgery ($p < 0.001$), wrong or delayed diagnosis ($p < 0.001$) and medication-related errors ($p < 0.001$) were the most important predictors for perceived incorrect treatment (table 3). In all, 12 out of 17 predictors were significantly associated with perceived incorrect treatment, which also included experiences related to doctor services, organisation and information. The regression model explained 39% of the variance.

DISCUSSION

There is a growing interest in involving patients in the evaluation of their safety.^{9–11} However, previous studies have not adequately documented the psychometric properties of the measurement instruments nor tested the reliability of patient-safety questions as quality indicators. The goal of this study was to document the development, psychometric testing and hospital-level reliability testing of the PRIH-I. The PRIH-I consists of 13 questions about patient-perceived incidents in hospitals and was tested in a national patient-experience survey. The missing-item rate and test–retest reliability for several items were poor, but hospital-level reliability was satisfactory for most items. The incident items contribute to a patient-reported incident index, with excellent data quality and hospital-level reliability.

The national patient-experience survey about hospitals was conducted among random samples of adult inpatients from all hospitals in Norway. This generic

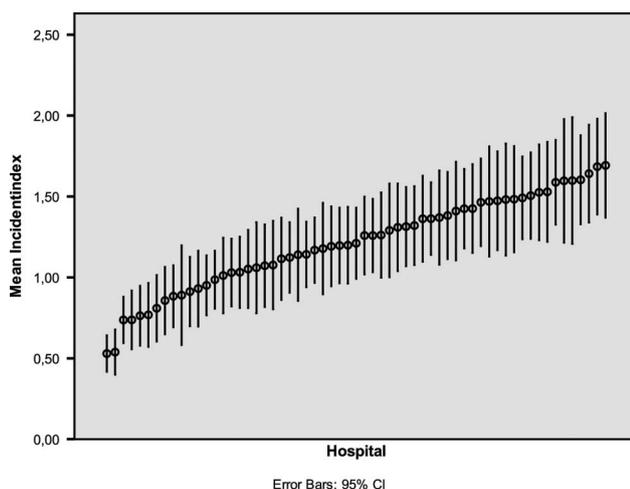


Figure 1 Mean score (95% CI) for the patient-incident index by hospital.

approach to patient-perceived incidents includes patients across all hospital departments for adults. We are only aware of one previous study that has used such a generic approach, but that study was limited to a single hospital.²² The present study was conducted among all hospitals in Norway, and the standardised methods ensure comparability between hospitals. This means that patient-perceived incidents can be aggregated and analysed at both the national and hospital levels, together with other quality indicators. The patient-incident index appears to be especially promising in a quality indicator setting, since it scored highly on all tests applied in this study.

However, both the development activities and empirical testing demonstrated that many hospital patients find patient-safety questions difficult to answer and/or of little relevance. For example, items about surgery are not relevant for patients who do not receive surgical interventions. This means that many patients have to respond to irrelevant questions by ticking ‘not applicable’ or the top box response option. The national survey showed that the amount of item-missing was large for some items, but this is a natural reflexion of relevance across the patient population. Combined with the declining response rate in the national survey, this meant that the goal was to develop and test a short incident instrument. The PRIH-I comprises 13 items, and the response rate in 2011 was similar to that of the last survey in 2006. This indicates that the inclusion of the PRIH-I in the national survey was acceptable to patients.

The importance of using lay language for patients reporting about their safety has been stressed previously.¹¹ We used cognitive interviews to test incident concepts and questions with patients. In psychometrics, items are considered to be empirical reflections of a latent construct, and multi-item scales are the standard approach. The reliability is normally lower for single items than for multi-item scales, which was also the case in the present study. All patient-experience scales had reliability estimates exceeding 0.7,¹⁹ but only one of 13 incident items met this criterion. The low level of test–retest reliability for many items shows the presence of substantial amounts of measurement error. This might be related to cognitive difficulties that remain with the safety questions, a lack of stability in the construct over time because of actual changes²³ or changes in patients’ internal values.²⁴ For instance, the question about staff forgetting to give important information to the patient only had a test–retest reliability on 0.16, implying that more than 80% of the variation between individuals is related to measurement error. The question is more general than the other questions, and the lack of specificity might have resulted in some patients evaluating different experiences in the test and retest questionnaires. More generally, it is important to note that the test–retest estimates are based on

Table 3 Multivariate linear regression model: global item about perceived incorrect treatment as a dependent variable and patient-reported experience and incident items as independent variables

	B	β	Significance (p value)
Did you suffer any unnecessary injury or unnecessary problem as a result of a surgical procedure or examination? Reference: some, large or very large extent			
Not at all or small extent	19.8	0.45	<0.001
Not relevant	19.7	0.42	<0.001
Did you get a wrong or delayed diagnosis in connection with your hospital stay? Reference: some, large or very large extent			
Not at all or small extent	15.4	0.33	<0.001
Not relevant	13.4	0.26	<0.001
Did you receive the wrong medication, or were you incorrectly medicated in any other way? Reference: some, large or very large extent			
Not at all or small extent	13.3	0.26	<0.001
Not relevant	12.5	0.22	<0.001
Doctor services	0.20	0.18	<0.001
Did you experience that the staff forgot to give you important information?	-4.7	-0.09	<0.001
General organization	1.6	0.07	<0.001
Information	0.07	0.07	<0.001
Did you experience any administrative mistakes during your hospital stay?	-2.4	-0.05	<0.001
Did you get an infection in connection with your hospital stay? Reference: yes			
No	2.3	0.04	<0.001
Had infection before admission	0.6	0.02	NS
Did you experience that important information about you had reached the right people?	0.8	0.03	<0.001
One main doctor	-0.6	-0.03	<0.01
Fixed group of nurses	-0.5	-0.03	<0.05
Nursing services	-0.0	-0.03	NS
Was an updated list of your medications reviewed with you when you were discharged from hospital? Reference: no			
Yes	-0.8	-0.02	NS
Not applicable	0.1	0.01	NS
When you were admitted, did you bring along an updated list of medications you were taking? Reference: no			
Yes	0.3	0.00	NS
Not applicable	0.1	0.00	NS
Did you experience insufficient hand hygiene (hand washing) among the staff? Reference: some, large or very large extent			
Not at all or small extent	0.5	0.00	NS
Not relevant	0.4	0.00	NS
Did you experience that the staff forgot to check your identity?	-0.1	0.00	NS

NS, not significant.

a rather small sample. In addition, the amount of variation on some variables is small, implying that relatively large effects might be caused by changes within a few individuals. It is not possible to determine the exact causes for individual changes across the set of dependent variables. However, the practical implications are to focus on the safety index (and possibly the most reliable single items) and to secure an adequate sample size at the hospital level. The patient-incident index consisting of 12 incident items had acceptable test-retest reliability, and hence constitutes a robust alternative to single items. Furthermore, when there were sufficiently large samples per hospital (n=200), all except one item reach the criterion of 0.7 for hospital-level reliability. The choice of outcome variables and sample size are important considerations when planning and designing a study of patient-perceived incidents. Based on this study, it seems advisable to recommend around 200

respondents per hospital and to use the incident index as a primary measure when comparing hospitals and monitoring changes over time.

The findings of this study demonstrate that patients have a broad perspective about hospital safety. The multivariate linear regression found that unnecessary injury following surgery, wrong or delayed diagnosis, and medication-related errors were the most important predictors for perceived incorrect treatment. However, organisation of hospitals, information and doctor services were strongly related to patient perception of overall incorrect treatment. This means that a full understanding of the patient perspective about hospital safety should include such dimensions. It also indicates that safety improvement work should be directed towards these more general quality issues. The percentage of patients perceiving incorrect treatment in Norwegian hospitals is stable over time. The percentage is somewhat lower than the estimate of 16%

of hospital stays in 2011 with at least one adverse event advent, estimated by the national safety campaign in Norway using the Global Trigger Tool (GTT) methodology. However, the quality of the GTT measurement in Norway has not been evaluated formally. The change over time in patient-based estimates and convergence with other sources is an important future research area.

Several previous studies have compared patient-safety incident reporting against other sources.^{25–29} While this is a generally relevant research area, it has little relevance for quality improvement efforts. Patient-reported safety and clinical patient safety are different concepts;¹⁰ the aim of the former is to improve our understanding and measure safety from the patient perspective, which is best evaluated by the patients themselves, while that of the latter is to assess safety in medical terms, which is best evaluated by medical professionals. Patient-reported incidents that are not confirmed by other sources are still quality problems: lack of convergence indicates communication problems with negative emotional consequences for patients, and hence poor patient-centeredness. Patient-centeredness is an important part of healthcare quality.^{1 2} Furthermore, it is not clear what to use as a gold standard for adverse events. For example, recent research on the inter-rater reliability of the Global Trigger Tool has shown that assessments by even experienced healthcare teams can differ substantially when there are adverse events in the same sample of records.⁷ All in all, the optimal solution seems to be to triangulate different methods and perspectives, and have a broad and complex view of safety measurement and improvement.

The timing of patient reports is considered an interesting research area.⁹ Many studies of patient experiences have found that making measurements at the time or shortly after discharge underestimates the amount of problems, because of social desirability bias and other sociopsychological factors.³⁰ Timing research needs to assess the appropriate balance between multiple factors: the need for hot reporting and safety improvement, sociopsychological pressures for social desirability bias, patient recovery and adequate mental distance to the incident, and the potential for recall bias. Patient-incident reporting at home after hospital discharge obviously increases the likelihood of underreporting of the most serious mistakes and deaths. To compensate for this, the patients' next-of-kin and other proxies should be able to respond on the patients' behalf. In addition, other safety reporting systems must be used to ensure full coverage of all safety incidents.

Limitations

The lack of a comprehensive theoretical model is a limitation of this study. The conceptual approach was further developed and refined as part of the project, but the study did not follow our standard

development and validation process.^{12–15} An internal reference group was used instead of an external group, and we only conducted cognitive interviews with patients, not indepth interviews. The inclusion of the PRIH-I within an existing patient-experience questionnaire gave rise to space restrictions. The existing patient-experience questionnaire was not developed and validated with a particular focus on patient-perceived safety, which means that broader dimensions of patient-perceived safety might be poorly represented. Furthermore, topics like falls, pressure sores and equipment failures are not included in the PRIH-I. Some of these incidents were excluded because they were difficult for patients to evaluate, but it nevertheless means that the PRIH-I does not have full coverage of all possible incident types. Future research could test other ways of formulating such questions to improve content validity.

The psychometric evaluation model differed from the standard approach by focusing on single items instead of multi-item scales. A single-item approach means that factor analysis and internal consistency reliability are not relevant. The revised assessment method included test-retest reliability and hospital-level reliability for all safety items, as well as the construction and testing of an incident index consisting of 12 incident items. We believe this to be a robust model, but stress the importance of using the Index and an appropriate sample size when the PRIH-I is applied at the hospital level. Furthermore, more research is needed on the PRIH-I, including its ability to measure changes over time, correlation with other quality and safety indicators at the hospital level, case-mix considerations, and research into how the PRIH-I can be used effectively in local quality improvement work. The last topic is especially important since one of the aims of national patient-experience surveys in Norway is to improve quality.

CONCLUSIONS

The PRIH-I consist of 13 questions about patient-perceived incidents in hospitals, and can be easily and cost-effectively included in a national patient-experience survey with only a small increase in respondent burden. Although the missing-item rate and test-retest reliability for several items were poor, hospital-level reliability was satisfactory for most items. The incident items contribute to a patient-reported incident index, with excellent data quality and hospital-level reliability.

Acknowledgements Thanks to Kari Aanjesen Dahle for being project leader for the development project. Thanks to Marit Skarpaas, Sinan Akbas, Ulla Benedicte Funder and Solveig Eggen for their contribution to the national data collection, and Tomislav Dimoski for developing the FS system, carrying out the technical aspects of the national survey and being project leader for data collection.

Contributors OB planned the paper together with KES, HHI and AKL, carried out the statistical analysis, and drafted the

paper. KES, HHI and AKL participated in the planning process, revised the draft critically and approved the final version.

Funding The study was financed by the Norwegian Knowledge Centre for the Health Services.

Competing interests None.

Ethics approval The Data Inspectorate and the Norwegian Ministry of Health and Care Services approved the survey.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Anonymous data can be made available upon request.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) 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 and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;260:1743–8.
- Arah OA, Westert GP, Hurst J, *et al.* A conceptual framework for the OECD Health Care Quality Indicators Project. *Int J Qual Health Care* 2006;18(Suppl 1):5–13.
- Crow R, Gage H, Hampson S, *et al.* The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002;6:1–244.
- Garratt AM, Solheim E, Danielsen K. *National and cross-national surveys of patient experiences: a structured review*. Oslo: Norwegian Knowledge Centre for the Health Services, Report 7-2008.
- Gerteis M, Edgman-Levitan S, Daley J, *et al.* *Through the patient's eyes: understanding and promoting patient-centered care*. San Francisco, CA: Jossey-Bass, 1993.
- Borzecki AM, Christiansen CL, Chew P, *et al.* Comparison of in-hospital versus 30-day mortality assessments for selected medical conditions. *Med Care* 2010;48:1117–21.
- Schildmeijer K, Nilsson L, Arestedt K, *et al.* Assessment of adverse events in medical care: lack of consistency between experienced teams using the global trigger tool. *BMJ Qual Saf* 2012;21:307–14.
- Fitzpatrick R, Davey C, Buxton MJ, *et al.* Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:i–iv, 1–74.
- Ward JK, Armitage G. Can patients report patient safety incidents in a hospital setting? A systematic review. *BMJ Qual Saf* 2012;21:685–99.
- Massó Guijarro P, Aranaz Andrés JM, Mira JJ, *et al.* Adverse events in hospitals: the patient's point of view. *Qual Saf Health Care* 2010;19:144–7.
- King A, Daniels J, Lim J, *et al.* Time to listen: a review of methods to solicit patient reports of adverse events. *Qual Saf Health Care* 2010;19:148–57.
- Iversen HH, Holmboe O, Bjertnaes OA. The Cancer Patient Experiences Questionnaire (CPEQ): reliability and construct validity following a national survey to assess hospital cancer care from the patient perspective. *BMJ Open* 2012;2:e001437.
- Skudal KE, Garratt AM, Eriksson B, *et al.* The Nordic Patient Experiences Questionnaire (NORPEQ): cross-national comparison of data quality, internal consistency and validity in four Nordic countries. *BMJ Open* 2012;2:e000864.
- Garratt AM, Bjertnaes OA, Holmboe O, *et al.* Parent experiences questionnaire for outpatient child and adolescent mental health services (PEQ-CAMHS Outpatients): reliability and validity following a national survey. *Child Adolesc Psychiatry Ment Health* 2011;5:18.
- Garratt AM, Bjertnaes OA, Barlinn J. Parent experiences of paediatric care (PEPC) questionnaire: reliability and validity following a national survey. *Acta Paediatr* 2007;96:246–52.
- Bjertnaes O, Skudal KE, Iversen HH. Classification of patients based on their evaluation of hospital outcomes: cluster analysis following a national survey in Norway. *BMC Health Serv Res* 2013;13:73.
- Pettersen KI, Veenstra M, Guldvog B, *et al.* The Patient Experiences Questionnaire: development, validity and reliability. *Int J Qual Health Care* 2004;16:453–63.
- Garratt AM, Helgeland J, Gulbrandsen P. Five-point scales outperform 10-point scales in a randomized comparison of item scaling for the Patient Experiences Questionnaire. *J Clin Epidemiol* 2011;64:200–7.
- Skudal KE, Holmboe O, Iversen H, *et al.* *Patient experiences with hospitals: national results in 2011 and changes from 2006*. Oslo, Norway: The Norwegian Knowledge Centre for the Health Services, PasOpp-report 3—2012.
- Terwee CB, Bot SD, de Boer MR, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
- Roland M, Elliott M, Lyrtzopoulos G, *et al.* Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. *BMJ* 2009;339:b3851.
- Agoritsas T, Bovier PA, Perneger TV. Patient reports of undesirable events during hospitalization. *J Gen Intern Med* 2005;20:922–8.
- Streiner DL, Norman GR. *Health measurement scales—a practical guide to their development and use*. 4th edn. Oxford, UK: Oxford University Press, 2008.
- Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med* 1999;48:1531–48.
- Weingart SN, Pagovich O, Sands DZ, *et al.* What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. *J Gen Intern Med* 2005;20:830–6.
- Friedman SM, Provan D, Moore S, *et al.* Errors, near misses and adverse events in the emergency department: what can patients tell us? *CJEM* 2008;10:421–7.
- Kaboli PJ, Glasgow JM, Jaipaul CK, *et al.* Identifying medication misadventures: poor agreement among medical record, physician, nurse, and patient reports. *Pharmacotherapy* 2010;30:529–38.
- Weissman JS, Schneider EC, Weingart SN. Comparing patient-reported hospital adverse events with medical record review: do patients know something that hospitals do not? *Ann Intern Med* 2008;149:100–8.
- van den Bemt PM, Egberts AC, Lenderink AW, *et al.* Adverse drug events in hospitalized patients. A comparison of doctors, nurses and patients as sources of reports. *Eur J Clin Pharmacol* 1999;55:155–8.
- Bjertnaes OA. The association between survey timing and patient-reported experiences with hospitals: results of a national postal survey. *BMC Med Res Methodol* 2012;12:13.