OXFORD

An analysis of complaints about hospital care in the Republic of Ireland

EMILY O'DOWD[1,2], SINÉAD LYDON [2,3], KATHRYN LAMBE[4], AKKE VELLINGA [5],
CHRIS RUDLAND[6], ELAINE AHERN[6], AOIFE HILTON[6], MARIE E. WARD [7], MARIA KANE[7,8],
TOM READER[9], ALEX GILLESPIE[9], DAVID VAUGHAN[10], DUBHFEASA SLATTERY[11], and
PAUL O'CONNOR[1,2]

[1]Discipline of General Practice, School of Medicine, National University of Ireland Galway, Galway H91TK33, Ireland
[2]Irish Centre for Applied Patient Safety and Simulation, School of Medicine, National University of Ireland Galway, Galway, Ireland
[3]School of Medicine, National University of Ireland Galway, Galway, Ireland
[4]Health Research Board, Dublin D02 H638, Ireland
[5]School of Public Health, Physiotherapy and Sport Science, University College Dublin D04 V1W8, Ireland
[6]National Complaints Governance and Learning Team, Health Service Executive, Catherine Street, Limerick V94 AY27, Ireland
[7]Quality and Safety Improvement Directorate, St James's Hospital, Dublin D08 NHY1, Ireland
[8]Centre for Innovative Human Systems, School of Psychology, Trinity College, The University of Dublin, Dublin D02 PN40, Ireland
[9]London School of Economics, London, UK
[10]Department of Paediatrics, Children's Health Ireland, Tallaght, Dublin, Ireland
[11]Health Service Executive, Dublin, Ireland

Address reprint requests to: Emily O'Dowd, Discipline of General Practice, National University of Ireland Galway, Galway H91TK33, Ireland. Tel: +00353 91495257, E-mail: Emily.odowd@nuigalway.ie

Abstract

Background: Patients and family members make complaints about their hospital care in order to express their dissatisfaction with the care received and prompt quality improvement. Increasingly, it is being understood that these complaints could serve as important data on how to improve care if analysed using a standardized tool. The use of the Healthcare Complaints Analysis Tool (HCAT) for this purpose has emerged internationally for quality and safety improvement. Previous work has identified hot spots (areas in care where harm occurs frequently) and blind spots (areas in care that are difficult for staff members to observe) from complaints analysis. This study aimed to (i) apply the HCAT to a sample of complaints about hospital care in the Republic of Ireland (RoI) to identify hot spots and blind spots in care and (ii) compare the findings of this analysis to a previously published study on hospital complaints in the UK.

Methods: A sample of complaints was taken from 16 hospitals in the RoI in Quarter 4 of 2019 ($n = 641$). These complaints were coded using the HCAT to classify complaints by domain, category, severity, stage of care and harm. Chi-squared tests were used to identify hot spots, and logistic regression was used to identify blind spots. The findings of this study were compared to a previously published UK study that used HCAT to identify hot spots and blind spots.

Results: Hot spots were identified in Irish hospital complaints while patients were receiving care on the ward, during initial examination and diagnosis, and while they were undergoing operations or procedures. This aligned with hot spots identified in the UK study. Blind spots were found for systemic problems, where patients experience multiple issues across their care.

Conclusions: Hot spots and blind spots for patient harm can be identified in hospital care using the HCAT analysis. These in turn could be used to inform improvement interventions, and direct stakeholders to areas that require urgent attention. This study also highlights the promise of the HCAT for use across different healthcare systems, with similar results emerging from the RoI and the UK.

Key words: quality improvement, hospitals, patient safety, statistical data analysis, patient satisfaction, patient-centred care

Introduction

Healthcare complaints are formal expressions of dissatisfaction with the care provided to patients [1]. They are typically communicated through a letter by a patient or their family [2]. Patient insights, as derived from complaints, are increasingly recognized as a useful source of data on how to improve care [3]. In order to promote an equitable, patient-centred health service, it is vital that we incorporate the patient perspective and extend the same level of respect to their input as is offered to healthcare providers and managers working within the system [3]. However, research suggests that although complaints about a unit can prompt changes within that unit, typically, the learning from complaints is not shared externally [4]. Accordingly, there is potential for complaints data to be used in the aggregate for quality and safety improvement in healthcare [1].

The Healthcare Complaints Analysis Tool (HCAT) has been developed to provide a systematic, reliable method of

accessing and analysing patient complaints about secondary care [5] and is intended to support the analysis of and learning from complaints and facilitate care improvement. The HCAT can be used to identify three distinct domains (clinical, management and relationship) within healthcare complaints, comprising seven categories (safety, quality, environment, institutional processes, listening, communication and respect and patient rights). The HCAT also categorizes the stage of care, harm and severity reported in the complaint [5].

The HCAT has been used to analyse patient complaints in a number of countries globally (all of which are considered developed economically) [5–9]. The tool has been found to remain reliable when applied in different health systems when used on claims for compensation and when translated [9–11]. It has demonstrated the potential to assist stakeholders in identifying areas in hospital care that require quality improvement [5]. Previous applications of the HCAT have identified 'hot spots' (i.e. points in care where there is a high occurrence of harm or near misses where harm could have occurred [12]) and 'blind spots' (i.e. areas in care that are not easily witnessed by staff members within an organization or that can be witnessed incorrectly by staff members [12]). It has been proposed that identifying these areas using the HCAT could help with the development of improvement interventions [7, 13].

The aim of our study was to apply the HCAT to patient complaints made about healthcare in the Republic of Ireland (RoI) and identify areas in care where quality improvement is required. This analysis also aims to provide a comparison with a previously published UK study that used the HCAT to analyse complaints in order to provide some context to the data from the RoI complaints [12].

Method
Design
This is a retrospective analysis of a database of complaints made by patients about hospitals in the RoI.

Setting
This study was conducted using complaints about hospital care in the public healthcare system in the RoI.

Sample
Individual hospitals were contacted by the Health Service Executive (HSE), National Complaints Governance and Learning Team, and asked to contribute their complaints from Quarter 4 of 2019 for analysis. The HSE is the health service in the RoI. Redacted complaints were obtained from 16 hospitals. For the purpose of this study, a 'complaint' is considered to be a formal expression of dissatisfaction with healthcare received, made by a patient, family member or other representatives to the hospital involved or the HSE. Claims for financial compensation and disciplinary complaints made to external governing bodies such as the Irish Medical Council were not included.

Ethical approval
This study received ethical approval from the NUI Galway Research Ethics committee (reference: 18 September 2017).

Procedure
The complaints were collated using Microsoft Excel. Each complaint was read thoroughly by one of four researchers (K.L., E.O.D., M.E.W. or M.K.), and then, the HCAT was used to categorize the complaints. A total of 25% of complaints were double coded by the second author to ensure inter-rater reliability was sufficient. It was decided to code 25% based on both the timeframe available for the project, as well as previous health research, which has determined 10% as the minimum recommended percentage for inter-rater reliability [14, 15]. The HCAT was applied to categorize individual issues that were reported within complaints in terms of the following:

- the domain of the issue(s)- clinical, management or relationship;
- problem categories of issue(s)—institutional processes, quality, respect and patient rights, communication, safety, environment or listening;
- stage of care at which the issue(s) occurred—admission, examination, ward, operation/procedure, discharge or unknown/other (e.g. when the stage was unclear);
- severity of issue(s)—1 (low) to 3 (high); and
- overall harm reported by the patient in the complaint—0 (no harm/harm not mentioned) to 5 (catastrophic harm).

Data preparation
Data were exported from Microsoft Excel into R statistical software [16] for analysis. Data were cleaned, and the 'harm' variable was transformed from a variable with six levels to a dichotomous variable. This was required for a subsequent part of the analysis as the sample size was too small for the statistical test to reliably run with a variable with six levels. This action has previously been taken in an analysis of general practice complaints [13] and was the approach recommended by an epidemiologist (A.V.). As a result, any complaint that did not report harm (i.e. harm = 0) was coded as 'No harm' and complaints that reported harm of any level (from minor to catastrophic) were coded as containing 'harm'. While only one level of harm is recorded per complaint, the other aspects of the HCAT are measured at the 'issue' level, that is, each problem within a complaint is considered separately.

Analysis
The analysis was made up of five distinct parts.

1. *Descriptive analysis.* Following coding, analyses of the features of the complaints were conducted by one researcher (E.O.D.). The characteristics of staff and complainants were captured where possible, and the frequencies of each HCAT aspect were also captured.
2. *Inter-rater reliability.* Complaints were initially coded using the HCAT by one researcher (K.L./E.O.D./ M.E.W./M.K.). Following this, 25% of complaints were double coded by a second researcher (either K.L. or E.O.D.). The agreement was calculated using Gwet's AC1 statistic. This is more robust than the traditionally used kappa statistic [17] and was considered appropriate for the purposes of this study as it has been used previously when employing the HCAT and could therefore be used for comparative purposes [5, 17].

3. *Identification of hot spots.* One of the core aspects of this analysis was to conduct an exploration of the hot spots that emerged. Hot spots are points in care where harm to patients is reported more frequently than would be expected [12]. There can also be hot spots for near misses—where a complaint is coded as high severity; however, no harm was reported, indicating a near-miss incident [12]. Hot spots were analysed using chi-squared tests (for trend), testing the relationship between reported harm and stage of care, following the R script written by Reader and Gillespie [12].

4. *Identification of blind spots.* Blind spots are the next aspect of this complaints analysis. Blind spots are points across care that staff members within a healthcare organization cannot easily measure or observe or indeed that cannot be observed at all [12]. However, patients often witness these issues. There are three types of blind spots. Blind spots for entry and exit into care are identified by computing the number of complaints issues that occur at Stages 1 (Admissions) and 5 (Discharge). Blind spots for errors of omission are defined within the HCAT as issues within the categories 'Quality', 'Listening' and 'Communication' [12]. Blind spots for systemic issues (i.e. where patients experience issues in care across multiple points in the care pathway) were analysed using logistic regression, with harm as the dependent variable, and number of issues and number of stages within a complaint as the independent variables in the model. No confounding variables were included in the model. Similar to the hot spot analysis, this used the code written by Reader and Gillespie [12].

5. *Narrative comparison with UK complaint data.* Data were extracted on key findings from the application of the HCAT to a sample of complaints about hospitals in the UK [12].

Results

Descriptive statistics

A total of 641 complaints, pertaining to Quarter 4 of 2019, were analysed. This represented 72% of all complaints (total $n = 896$) received by Irish hospitals in this period and allowed for a 2% margin of error with a 95% confidence level. This means that we can be confident to the 95% level that across all of the complaints received by Irish hospitals, our values will fall within 2% of the values in this paper. Within the sample of complaints, there were a total of 1308 unique issues identified. There was a mean of 2.05 issues per complaint (SD = 1.23) and a range of 1–6 issues. A total of 16 hospitals provided data for this study, with an average of 40 complaints provided by each hospital (range = 6–146).

Harm

Almost half of the complaints in this sample reported no harm ($n = 308$, 48%). However, 12 (2%) of the complaints recorded catastrophic harm, i.e. permanent injury or death. A summary of the harm within complaints can be found in Table 1.

Categories

All HCAT categories were represented in the complaints. Institutional processes problems made up 30% of all

Table 1 Descriptive statistics from present analysis and comparison to UK data

HCAT domains and categories	Present study, Ireland N (%)	Gillespie and Reader, UK N (%)
Total complaints	641	1110
Total complaints issues	1308	2047
Clinical		
Quality	189 (14%)	328 (16%)
Safety	160 (12%)	340 (16%)
Management		
Environment	115 (9%)	192 (9%)
Institutional processes	390 (30%)	523 (25%)
Relationships		
Listening	92 (7%)	133 (6%)
Communication	180 (14%)	259 (12%)
Respect and patient rights	182 (14%)	274 (13%)
Severity		
Low	292 (22%)	558 (27%)
Medium	726 (56%)	1030 (50%)
High	287 (22%)	486 (23%)
Uncoded	3 (0%)	NA
Stage of care		
1. Admissions	322 (25%)	353 (17%)
2. Examination/ diagnosis	233 (18%)	443 (21%)
3. Care on the ward	370 (28%)	465 (22%)
4. Operation and procedures	78 (6%)	227 (11%)
5. Discharge	68 (5%)	310 (15%)
6. Other	171 (13%)	135 (7%)
Multiple stages	62 (5%)	75 (4%)
Not clear/uncoded	4 (0%)	66 (3%)
Harm	(Per complaint)	(Per complaint)
0—No harm reported	308 (48%)	409 (37%)
1—Minimal	112 (17%)	48 (4%)
2—Minor	114 (18%)	248 (22%)
3—Moderate	58 (9%)	152 (14%)
4—Major	28 (4%)	163 (15%)
5—Catastrophic	12 (2%)	90 (8%)
Unclear/not coded	9 (1%)	0 (0%)
Hot spots	Examination/ diagnosis Care on ward Operation/ procedure	Examination/ diagnosis Care on ward Operation/ procedure
Blind spots	Entry/exit Systemic problems Errors of omission	Entry/exit Systemic problems Errors of omission

issues ($n = 390$), with quality ($n = 189$, 14%), communication ($n = 180$, 14%) and respect and patient rights ($n = 182$, 14%) representing the next most frequently mentioned categories. Table 1 details the distribution across the categories.

Severity

The majority of issues were classified as medium severity ($n = 726$, 56%). Table 1 presents a breakdown of all severity ratings.
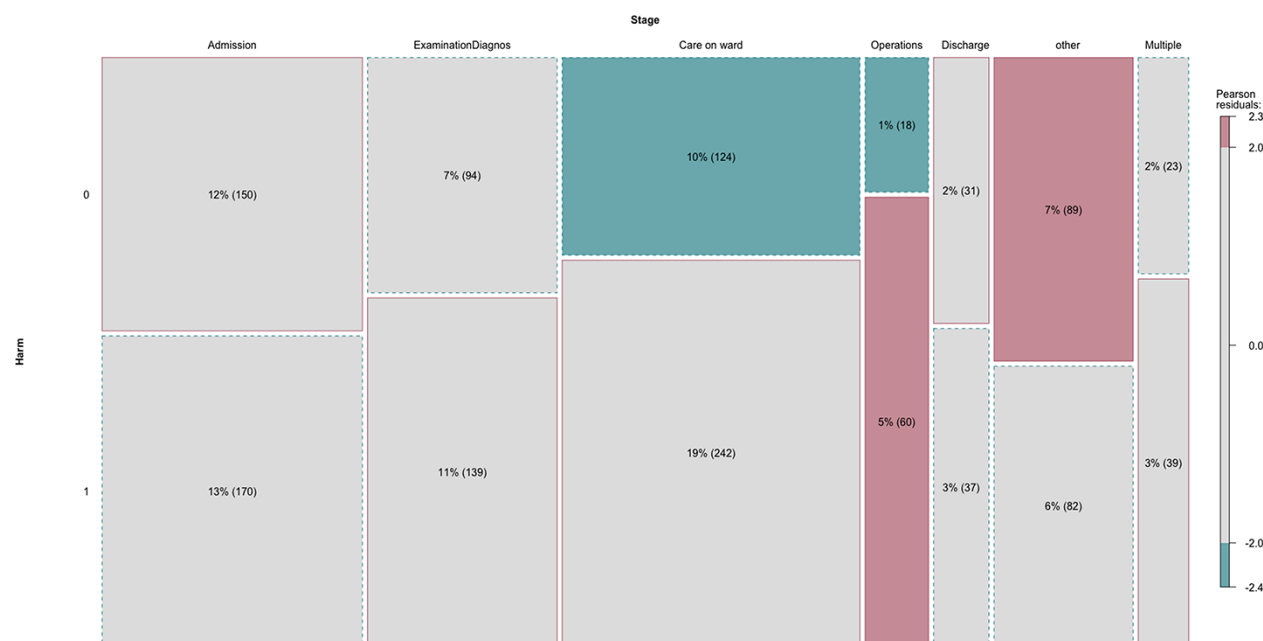
Figure 1 Hot spots for harm by stages of care.

Stage of care

The highest proportion of complaints issues ($n = 370$, 28%) referred to problems while the patients were receiving care on the ward (Stage 3 of the HCAT). This was closely followed by Stage 1, Admissions ($n = 322$, 25%). Table 1 presents a breakdown of the stages of care.

Inter-rater reliability

Gwet's AC1 across all aspects of the HCAT was 0.93 (CI = 0.92–0.94, $P < 0.05$). This was consistent with the findings of the original study by Reader and Gillespie in the UK using the HCAT, which had reliability ranging from 0.69 to 0.91 across categories [5].

Hot spots

The relationship between harm and stage of care as calculated using the chi-squared test was significant (X^2 $(6, n = 1308) = 32.27$, $P < 0.05$). This indicated that there were hot spots for harm, which were identified at three stages—care on the ward, examination and diagnosis and operations and procedures. There was also a hot spot when issues contained multiple stages of care. Figure 1 represents visually the hot spots at each of the stages of care. Blue boxes/a blue border indicates that the observed number of complaints issues at that point were fewer than expected, with red boxes/red border indicating that the observed number was more than expected. The colour of a box indicates the magnitude of the difference between the expected numbers of complaint issues per stage and the actual observed number of issues per stage. For example, a red box has more than expected complaints issues at that stage of care, and the value is farther from the expected value than a grey box with a red border. A blue box has fewer than expected complaints issues at that stage, and this number is further from the expected value than a grey box with a dotted blue border.

Table 2 Full logistic regression model for blind spots for harm

Coefficients	Estimate	Std error	Z-value	P
Intercept	−1.12	0.21	−5.34	<0.005*
Number of issues	0.36	0.09	3.92	<0.005*
Number of stages	0.35	0.18	1.96	0.0502

*Significant at a $P < 0.005$ level.

Blind spots

Blind spots were identified at admissions and discharge and implicitly for errors of omission. Finally, a blind spot for systemic problems across the patient care pathway was identified using logistic regression. The model indicated that as the number of issues and stages per complaint increased, so too did the likelihood that harm would be reported within a complaint. Table 2 presents the model.

Comparison with UK complaints analysis

Table 1 presents a comparison with the UK study of complaints in hospitals. The distribution of problems across categories, stages of care and severity levels were broadly similar across the two countries. The same hot spots and blind spots also emerged from the RoI and UK analyses.

Discussion
Statement of principal findings

Our analysis of complaints about hospital care using the HCAT allowed us to identify areas where harm was frequently reported by patients or family members (hot spots), and points in care where patients reported issues that staff could not see (blind spots). Hot spots occurred when patients were receiving care on the ward, during the examination and while undergoing a procedure. Blind spots included the boundaries of care and when patients experienced problems across the entire care

pathway. Such data provide clear direction for where interventions are required to improve patient safety and quality of care. It also highlighted that the findings of the HCAT analysis of complaints about hospital care in the RoI are broadly comparable to those in the UK.

Strengths and limitations

This study has several strengths and limitations. Strengths include that the study adds to the growing literature on the benefits of the HCAT for identifying issues within patient complaints [7, 12]. Second, this study reported high reliability for the HCAT in an Irish context, supporting existing data on the reliability of this tool [5, 9]. Third, this study involved multiple researchers and coders applying the HCAT, which reflects best practices [18]. Finally, the sample size achieved by this study was sufficient to provide a margin of error of only 2%, with a 95% confidence level [19].

Despite these strengths, there were also a number of limitations. Firstly, the sample size, while sufficient, restricted the complexity of analyses that were possible with some variables. Sampling complaints across a longer time frame could avoid this issue in future. The onset of the COVID-19 pandemic also inhibited the collection of complaints for this study, with many health service staff being redeployed to tackle the emerging public health crisis [20]. Further, a limitation of complaints research in general is that they represent only a small subsample of patients, who are willing and able to complain. Others may not be in a position to complain due to a lack of agency or knowledge or indeed out of fear of negative consequences to their care [21]. Finally, this study considered pre-COVID complaints only, and only complaints within an Irish and UK setting. Further research should explore how complaints may have changed since the beginning of the pandemic, and how the specific context of other countries may impact the categories that emerge from complaints.

Interpretation within the context of wider literature

The hot spots and blind spots that emerged from the complaints about care in the RoI were similar to those that were reported in the UK [12]. To date, no other study has reported on hot spots and blind spots; however, these metrics could be vital for future quality improvement work. Such analyses could support a focus on areas in care that may require quality improvement. Complaints are an existing source of data in many healthcare settings and could be used to improve care where it is needed the most urgently [22]. Identifying areas such as these hot spots where harm is likely to occur is important for quality improvement and would allow for a prospective approach to improvement, rather than the retrospective approach taken by many other metrics [23]. Examples of this could include sharing the findings of a complaints analysis with stakeholders to develop and apply interventions. The UK and RoI health systems are relatively similar in structure [24, 25], which may explain why there were similar hot spots and blind spots within the complaints in both jurisdictions. It would be interesting to conduct similar analyses in other healthcare systems to determine whether the same, or different, hot spots and blind spots emerge.

While this study analysed complaints about hospitals in the RoI, the proportions of complaints issues that were assigned

to each HCAT category were broadly similar to the UK [12]. In both studies, the 'institutional processes' category emerged most frequently. This seems to hold across other countries that have applied the HCAT [6–8] and highlights that complaints from individual patients can be indicative of system-level issues [22]. Understanding how patient complaints can shed light on management issues within a hospital system will be vital for providing evidence for interventions at this structural level [3, 12].

Harm was reported in over half of the complaints analysed from hospitals in the RoI. This is a key finding and links directly to the widely researched issue of iatrogenic harm in healthcare [26]. Identifying harm to patients using complaints could support staff-based measures of harm and add to our understanding of how this might be manifested [3, 23, 27]. The levels of harm reported in the RoI complaints were broadly similar to the UK [12]. It is also interesting that a large number of complaints reported no harm, which could be indicative of near-miss incidents that are perceived by patients and may not be captured in other staff-based metrics such as incident reports [3]. Future work should explore how complaints could be linked with other metrics to improve quality and safety in hospital care [28].

Implications for policy, practice and research

This study has implications for policy, practice and research into patient complaints. It highlights the reliability of the HCAT, supporting the use of the tool to analyse patient complaints. The HCAT, and the taxonomy from which it was developed, is increasingly being used internationally to classify patient complaints in different settings (e.g. tertiary hospitals, maternity care and general practice [6, 8, 13, 29]). It has also been translated into Danish, and reliability has been found to hold in the translated version [10], as well as when used on compensation claims in Denmark [11]. It also exemplifies the benefits of analysing existing complaints data in a structured manner, in that any recommendations for improvements are evidence-based and reliable [22, 30]. However, complaints are only one data source. It is recommended that future research should consider how complaints can be linked with other methods of measuring and monitoring quality and safety in order to obtain a broader understanding of issues in healthcare than can be achieved by considering these data in isolation [28]. It could also be useful to examine compensation claims in the Irish healthcare system in a future study in a similar manner to the work emerging from Denmark [11].

Related to the previous point, there is arguably little point in analysing complaints data if it is not then used to bring about improvement. These data should be used not only to identify areas for improvement but also to support the design and evaluation of interventions to improve patient safety and quality of care [31]. It is suggested that the analysis of the hot spots and blind spots may be of particular use in identifying the focus of potential interventions and that a co-design approach with stakeholders could be used to achieve this, as with other areas of patient safety [32]. In a global healthcare system with entrenched resourcing and funding issues [33], using a tool such as the HCAT to prioritize interventions for patient safety and quality improvement has the potential to be particularly useful for implementing organizational change.

Conclusions

The HCAT is a reliable tool when applied to the analysis of complaints about hospital care in the RoI. This study highlights the promise of the HCAT as a tool for the analysis of complaints across different healthcare systems, with similar results emerging in the RoI and the UK. The value of the HCAT for identifying hot spots and blind spots as areas for improvement is also supported by this study. However, for the analysis of complaints to be more than an academic exercise, it is important that moving forward, findings are used to drive the implementation of improvement interventions to address the issues identified.

Author contributions

All authors named on the paper have contributed sufficiently to the design and preparation of this study and manuscript. E.O.D. analysed complaints using the HCAT, conducted data analysis, completed data agreement and wrote the first draft of the paper. S.L. designed the study and drafted the paper. K.L. analysed complaints using the HCAT, completed the data agreement and reviewed and re-drafted the paper. A.V. contributed to the design of the study and statistical analysis and drafted the paper. C.R. contributed to the design of the study, aided in data collection and drafted the paper. E.A. contributed to data collection and anonymization and drafted the paper. A.H. contributed to the design of the study and aided in data collection. M.E.W. aided in data collection and coding of complaints and drafted the paper. M.K. aided in data collection and coding of complaints and drafted the paper. T.R. contributed to the design of the study and aided in the drafting of the paper. A.G. contributed to the design of the study and aided in the drafting of the paper. D.V. was involved in the design of the study and read the paper before finalizing. D.S. was involved in the design of the study and read the paper before finalizing. P.O.C. was the principal investigator on the grant, designed the study, drafted the paper, aided in data analysis and collaborated with data controllers to access data.

Ethics and other permissions

This study received ethical approval from the NUI Galway Research Ethics committee (reference: 18 Sept 17).

Data sharing statement

Data will be made available by the authors on reasonable request.

References

1. Reader TW, Gillespie A, Roberts J. Patient complaints in healthcare systems: a systematic review and coding taxonomy. *BMJ Qual Saf* 2014;**23**:678–89.
2. Bark P, Vincent C, Jones A *et al.* Clinical complaints: a means of improving quality of care. *Qual Saf Health Care* 1994;**3**:123–32.
3. O'Hara JK, Aase K, Waring J. Scaffolding our systems? Patients and families 'reaching in' as a source of healthcare resilience. *BMJ Qual Saf* 2019;**28**:3–6.
4. Liu JJ, Rotteau L, Bell CM *et al.* Putting out fires: a qualitative study exploring the use of patient complaints to drive improvement at three academic hospitals. *BMJ Qual Saf* 2019;**28**:894–900.
5. Gillespie A, Reader TW. The Healthcare Complaints Analysis Tool: development and reliability testing of a method for service monitoring and organisational learning. *BMJ Qual Saf* 2016;**25**:937–46.
6. Nowotny BM, Loh E, Davies-Tuck M *et al.* Identifying quality improvement opportunities using patient complaints: feasibility of using a complaints taxonomy in a metropolitan maternity service. *J Patient Saf Risk Manag* 2019;**24**:184–95.
7. Jerng J-S, Huang S-F, Yu H-Y *et al.* Comparison of complaints to the intensive care units and those to the general wards: an analysis using the healthcare complaint analysis tool in an academic medical center in Taiwan. *Crit Care* 2018;**22**:1–12.
8. van den Berg PF, Yakar D, Glaudemans AW *et al.* Patient complaints in radiology: 9-year experience at a European tertiary care center. *Eur Radiol* 2019;**29**:5395–402.
9. Bogh SB, Kerring JH, Jakobsen KP *et al.* Healthcare Complaints Analysis Tool: reliability testing on a sample of Danish patient compensation claims. *BMJ Open* 2019;**9**.
10. Morsø L, Walløe S, Birkeland S *et al.* Quantification of complaint and compensation cases by introducing a Danish translated and cross-cultural adapted edition of the Healthcare Complaints Analysis Tool. *Risk Manag Healthc Policy* 2021;**14**:1319–26.
11. Morsø L, Birkeland S, Walløe S *et al.* Compensation claims in Danish emergency care: identifying hot spots and blind spots in the quality of care. *Jt Comm J Qual Patient Saf* 2022. ahead of print.
12. Gillespie A, Reader TW. Patient-centered insights: using health care complaints to reveal hot spots and blind spots in quality and safety. *Milbank Q* 2018;**96**:530–67.
13. O'Dowd E, Lydon S, Lambe K *et al.* Identifying hot spots for harm and blind spots across the care pathway from patient complaints about general practice. *Fam Pract* 2021;cmab109.
14. Eagan BR, Rogers B, Serlin R *et al.* Can we rely on IRR? Testing the assumptions of inter-rater reliability. In: *International Conference on Computer Supported Collaborative Learning*, 2017.
15. Savona A, Ferralis L, Saffioti M *et al.* Evaluation of intra- and inter-rater reliability and concurrent validity of the Italian version of the Jebsen–Taylor Hand Function Test in adults with rheumatoid arthritis. *Hand Ther* 2019;**24**:48–54.
16. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing 2020.
17. Wongpakaran N, Wongpakaran T, Wedding D *et al.* A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol* 2013;**13**:61.
18. Boyer KK, Verma R. Multiple raters in survey-based operations management research: a review and tutorial. *Prod Oper Manag* 2000;**9**:128–40.
19. Kotrlik J, Higgins C. Organizational research: determining appropriate sample size in survey research appropriate sample size in survey research. *Inf Technol Learn Perform J* 2001;**19**:43.
20. Health Service Executive. Policy and procedure redeployment of staff during COVID-19 infection. 2020.
21. O'Dowd E, Lydon S, O'Connor P. A multi-perspective exploration of the understanding of patient complaints and their potential for patient safety improvement in general practice. *Eur J Gen Pract* 2021;**27**:35–44.
22. van Dael J, Reader TW, Gillespie A *et al.* Learning from complaints in healthcare: a realist review of academic literature, policy evidence and front-line insights. *BMJ Qual Saf* 2020;**29**:684–95.

23. Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety: drawing together academic evidence and practical experience to produce a framework for safety measurement and monitoring. 2013, April 2013.

24. National Health Service. 2019. *About NHS Hospital Services*. https://www.nhs.uk/nhs-services/hospitals/about-nhs-hospital-services/.

25. Health Service Executive. 2021. *Hospital Charges*. https://www.hse.ie/eng/about/who/acute-hospitals-division/patient-care/hospital-charges/.

26. Panagioti M, Khan K, Keers RN *et al*. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. *BMJ* 2019;**366**:l4185.

27. Ward JK, Armitage G. Can patients report patient safety incidents in a hospital setting? A systematic review. *BMJ Qual Saf* 2012;**21**:685–99.

28. Van Dael J, Gillespie A, Reader T *et al*. Getting the whole story: integrating patient complaints and staff reports of unsafe care. *J Health Serv Res Pol* 2021;**27**:41–9.

29. Wallace E, Cronin S, Murphy N *et al*. Characterising patient complaints in out-of-hours general practice: a retrospective cohort study in Ireland. *Brit J Gen Pract* 2018;**68**:e860–e8.

30. Gillespie A, Reader TW. Identifying and encouraging high-quality healthcare: an analysis of the content and aims of patient letters of compliment. *BMJ Qual Saf* 2020;bmjqs-2019-010077.

31. Sauro K, Ghali WA, Stelfox HT. Measuring safety of healthcare: an exercise in futility? *BMJ Qual Saf* 2020;**29**:341.

32. Lambe K, Lydon S, McSharry J *et al*. Identifying interventions to improve hand hygiene compliance in the intensive care unit through co-design with stakeholders. *HRB Open Res* 2021;**4**.

33. Baghbanian A, Torkfar G. Economics and resourcing of complex healthcare systems. *Austral Health Rev* 2012;**36**:394–400.