Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis

Maria Panagioti,1 Kanza Khan,1 Richard N Keers,2 Aseel Abuzour,2 Denham Phipps,2 Evangelos Kontopantelis,1 Peter Bower,1 Stephen Campbell,1 Razaan Haneef,3 Anthony J Avery,4 Darren M Ashcroft1

ABSTRACT

OBJECTIVE
To systematically quantify the prevalence, severity, and nature of preventable patient harm across a range of medical settings globally.

DESIGN
Systematic review and meta-analysis.

DATA SOURCES
Medline, PubMed, PsycINFO, Cinahl and Embase, WHOLIS, Google Scholar, and SIGLE from January 2000 to January 2019. The reference lists of eligible studies and other relevant systematic reviews were also searched.

REVIEW METHODS
Observational studies reporting preventable patient harm in medical care. The core outcomes were the prevalence, severity, and types of preventable patient harm reported as percentages and their 95% confidence intervals. Data extraction and critical appraisal were undertaken by two reviewers working independently. Random effects meta-analysis was employed followed by univariable and multivariable meta regression. Heterogeneity was quantified by using the I2 statistic, and publication bias was evaluated.

RESULTS
Of the 7313 records identified, 70 studies involving 337 025 patients were included in the meta-analysis. The pooled prevalence for preventable patient harm was 6% (95% confidence interval 5% to 7%). A pooled proportion of 12% (9% to 15%) of preventable patient harm was severe or led to death. Incidents related to drugs (25%, 95% confidence interval 16% to 34%) and other treatments (24%, 21% to 30%) accounted for the largest proportion of preventable patient harm. Compared with general hospitals (where most evidence originated), preventable patient harm was more prevalent in advanced specialties (intensive care or surgery; regression coefficient b=0.07, 95% confidence interval 0.04 to 0.10).

CONCLUSIONS
Around one in 20 patients are exposed to preventable harm in medical care. Although a focus on preventable patient harm has been encouraged by the international patient safety policy agenda, there are limited quality improvement practices specifically targeting incidents of preventable patient harm rather than overall patient harm (preventable and non-preventable). Developing and implementing evidence-based mitigation strategies specifically targeting preventable patient harm could lead to major service quality improvements in medical care which could also be more cost effective.

Introduction
Patient harm during healthcare is a leading cause of morbidity and mortality internationally.1 2 The World Health Organization defines patient harm as “an incident that results in harm to a patient such as impairment of structure or function of the body and/or any deleterious effect arising there from or associated with plans or actions taken during the provision of healthcare, rather than an underlying disease or injury, and may be physical, social or psychological (eg, disease, injury, suffering, disability and death).”3 The health burden and patient experiencing healthcare-related patient harm has been reported to be comparable to chronic diseases such as multiple sclerosis and cervical cancer in developed countries, and tuberculosis and malaria in developing countries.4 5 Harmful patient incidents are also a major financial burden for healthcare systems across the globe. It is estimated that 10-15% of healthcare expenditure is consumed by the direct sequelae of healthcare-related patient harm.6 7 Early detection and prevention of patient harm in healthcare is an international policy priority.8 In principle, zero harm would be the ideal goal. However, this goal is not feasible because some harms cannot be avoided in clinical practice. For example, some adverse drug reactions which occur in the absence of any error in the prescription process and without the possibility...
of detection are less likely to be preventable. In recent years, the recognition that a proportion of patient harm is not preventable has increased attention to the notion of preventable patient harm. Most studies classify patient harm as preventable if it occurs as a result of an identifiable modifiable cause, and its future recurrence can be avoided by reasonable adaptation to a process, or adherence to guidelines, although universal consensus has not been established. Key sources of preventable patient harm could include the actions of healthcare professionals (errors of omission or commission), healthcare system failures, or involve a combination of errors made by individuals, system failures, and patient characteristics. Strengthening the focus on preventable patient harm has the potential to lead to greater tangible clinical benefits and improved translation of patient safety research findings into clinical practice. Patient safety improvement strategies underpinned by better understanding of the nature of preventable patient harm have greater prospects of efficiency (because they are more specific) and implementation (because clinicians can readily recognise their value).

There are several systematic reviews on overall patient harm across different medical settings, but none of these have focused on preventable patient harm. We undertook a systematic review and meta-analysis to estimate the prevalence of preventable patient harm across medical settings including hospitals, various specialties, and in primary care. We also examined the severity and most commonly occurring types of preventable patient harm. The secondary outcomes were the severity and types of preventable patient harm. In accordance to the reporting format of the eligible studies, severity of preventable patient harm was classified into mild, moderate, and severe. Key types of preventable harm were drug-related, diagnostic, medical procedure-related, and healthcare-acquired infections (definitions are presented in eTable 1).

We excluded the following: studies reporting data on harm but not on preventable patient harm; studies with an exclusive focus on a specific type of harm only (only drug-related harm) or a specific severity level of harm only (incidents which only resulted in readmissions or extended length of stay) because such estimates would differ from estimates based on any type or any severity level of preventable patient harm; and studies focused on specific patient populations (eg, patients with a particular disease) because such estimates could differ from estimates in the general population.

**Searches**

We searched five electronic bibliographic databases from January 2000 to 27 January 2019: Medline, Cinahl, Embase, Pubmed, and PsycINFO. We supplemented these searches by screening grey literature sources including three databases (WHOLIS, Google Scholar, SIGLE), relevant reports, and conference abstracts. We also screened existing systematic reviews and checked the reference lists of eligible studies. The search strategy is available in eTable 3.

**Study selection and extraction**

We exported the results of the searches to Endnote X8 and removed duplicates. We completed screening in two stages. Initially, the titles and abstracts of the studies were screened for eligibility. Afterwards, the full texts of studies initially assessed as relevant for the review were retrieved and checked against our inclusion or exclusion criteria. We devised a data extraction spreadsheet, after being piloted, to extract descriptive data on key study characteristics (eg, number and age of participants, research design, data collection, assessment of preventability) and quantitative outcomes (prevalence, types, and severity of preventable patient harm). Two independent researchers (KK and MP) performed the screening and data extraction with disagreements resolved by discussion within the wider team (AA, DA, RH, RK). The inter-rater reliability was excellent (kappa=0.88 and 0.90).

**Risk of bias assessment**

We evaluated the risk of bias in the studies by using an adapted form of the Newcastle Ottawa scale for cross sectional and cohort studies. This assessed the representativeness of the sample, sample size, response rate, ascertainment of the exposure, control of confounding variables, assessment of preventability, and appropriate statistical analysis, which provided a score ranging from 0 (lowest grade) to 9 (highest grade). A higher grade indicated a lower risk of bias.
For our analyses, studies scoring 7 or above were considered as low risk, whereas studies scoring below 7 were considered as high risk.

Analyses

Our primary outcome was the prevalence of preventable patient harm expressed as the proportion of patients with at least one preventable patient harmful incident and stratified according to different medical services. We also calculated and reported the median prevalence of preventable patient harm and interquartile ranges across all medical care settings. Our secondary outcomes were the severity and types of preventable patient harm expressed as proportions of the total number of preventable patient harmful incidents. We pooled all data in Stata 15 by using the metaprop command. To improve the meaning and interpretation of our findings in relation to the prevalence, severity, and common types of preventable patient harm, we also present data on the prevalence, severity, and common types of overall harm (preventable and non-preventable) by using the same pool of studies in all analyses.

We conducted univariable and multivariable meta regression to test the influence of study level moderators on the prevalence of preventable patient harm using the metareg command. Consistent with the recommendations of Thompson and Higgins, eight prespecified study level moderators were hypothesised to have an effect on the prevalence of preventable patient harm (medical setting, population, research design, assessment method of harm, assessment of preventability, sample size, risk of bias, WHO region). Moderators were selected and coded following consensus procedures and each moderator value was based on a minimum of eight studies. Covariates meeting our significance criterion (P<0.10) were entered into a multivariable meta regression model. The P<0.10 threshold was conservative, to avoid prematurely discounting potentially important explanatory variables. Because proportions were often expected to be small, we used Freeman-Tukey Double Arcsine transformation to stabilise the variances and then performed a random effects meta-analysis implementing the DerSimonian-Laird method.

Random effects models were used in all analyses because they are more conservative and have better properties in the presence of heterogeneity. Heterogeneity was quantified by using the I² statistic. Conventionally, I² values of 25%, 50%, and 75% indicate low, moderate, and high heterogeneity, respectively. We inspected the symmetry of the funnel plots and used Egger’s test to examine for publication bias. Funnel plots were constructed using the metafunnel command, and the Egger test was computed using the metabias command.

Patient and public involvement

Two patient partners, who were members of our research advisory panel, were involved in the development of our research questions and in selecting the outcome measures of this study. The two patients also provided critical feedback to the protocol of the systematic review and advised on the interpretation and dissemination of results.

Results

The searches yielded 7313 citations. After we removed duplicates and reviewed the titles and abstracts, 6522 articles were excluded. Of the remaining 307 stories, 241 were excluded after reviewing the full article. A total of 66 studies reporting 70 independent samples were included in the review. Figure 1 shows the study flow for the selection process.

Descriptive characteristics

This review is based on a pooled sample of 337,025 patients, 28,150 of whom experienced harmful incidents and 15,419 experienced preventable harmful incidents. A total of 47,148 harmful incidents were identified in the pooled sample, 25,977 (55%) of which were preventable. The sample sizes ranged widely across studies (median 1440 patients, range 128-96,047). Thirty-three studies (47%) were conducted in the US, 27 (39%) in Europe, and 10 (14%) elsewhere. The most common study design was retrospective or cross-sectional (n=50; 71%) followed by prospective (20; 29%). Fifty-three studies (76%) reviewed the medical charts of patients to detect harm, whereas 17 studies (24%) monitored patients over time or were based on self reports (eg, interviews with patients). All included studies assess the preventability of patient harm by using consensus procedures between two or more trained reviewers (physicians or teams of physicians and nurses). Fifty studies (71%) used a standardised Likert scale to facilitate the consensus decisions for the preventability of patient harm among the reviewers (harmful incidents assigned a score of four out of six and over were considered preventable). The remaining 20 studies (29%) used implicit agreed criteria to reach consensus regarding the preventability of patient harm among the reviewers. Most studies were conducted in general hospitals involving patients from a range of specialties (45 studies; 64%). Twelve studies (17%) were conducted in advanced care specialties (intensive care 6 studies; surgery 6 studies), six studies (8%) in emergency department, four in obstetrics (6%), and three in primary care (4%). Except for six studies (9%), which were based on children and adolescents, and five studies on older adults (7%), the remaining 59 studies (84%) were mainly based on adults. Further details of the descriptive characteristics of the included studies are available in eTable 2.

All 70 studies reported data on the prevalence of preventable patient harm and overall patient harm. One third of the studies (20 studies, 29%) reported data on the severity of preventable patient harm. Forty-three studies (60%) reported proportions of at least two of the following six types of preventable patient harm: drug management, non-drug therapeutic management, diagnosis, invasive medical procedures, and dissemination of results.
surgical procedures, and infections acquired during healthcare.

Risk of bias results
The Newcastle Ottawa scores for the studies ranged from three to nine (maximum 9, a higher score indicating a lower risk of bias). Twenty-nine studies (41%) scored eight or above and were considered to be at low risk of bias (see full assessment in eTable 3).

Meta-analysis of the prevalence of preventable patient harm stratified by medical settings
Table 1 shows that the pooled prevalence of preventable patient harm was 6% (95% confidence interval 5% to 7%, $I^2=99$%) and the median prevalence was 5% (interquartile range 3-9%). In comparison, the pooled prevalence of overall harm (preventable and non-preventable) was 12% (95% confidence interval 9% to 14%, $I^2=99$%; table 1) and the median was 10% (interquartile range 7-15%). The highest pooled prevalence estimate of preventable patient harm was reported in intensive care (18%, 95% confidence interval 12% to 26%, $I^2=96$%) and surgery (10%, 7% to 13%, $I^2=97$%) and the lowest in obstetrics (2%, 0% to 4%, $I^2=95$%). Figure 2 presents the forest plot of the prevalence of preventable patient harm across medical care settings.

Meta-analysis of the severity and types of preventable patient harm
Table 1 shows the pooled proportions of the severity and types of preventable patient harm. The pooled proportion of mild harm was 49% (95% confidence interval 43% to 56%, $I^2=97$%), moderate harm was 36% (31% to 42%, $I^2=96$%), and severe harm was 12% (9% to 15%, $I^2=94$%).

Drug management incidents (25%, 95% confidence interval 16% to 34%, $I^2=98$%), and other therapeutic management incidents (24%, 21% to 30%, $I^2=98$%), accounted for the highest proportion of preventable patient harm followed by incidents related to surgical procedures (23%, 9% to 38%, $I^2=98$%), healthcare infections (16%, 11% to 22%, $I^2=98$%), and diagnosis (16%, 11% to 21%, $I^2=98$%).

Meta-regressions exploring the variance in the prevalence of preventable patient harm
Table 1 shows the pooled proportions of the severity and types of preventable patient harm. The pooled proportion of mild harm was 49% (95% confidence interval 43% to 56%, $I^2=97$%), moderate harm was 36% (31% to 42%, $I^2=96$%), and severe harm was 12% (9% to 15%, $I^2=94$%).

Drug management incidents (25%, 95% confidence interval 16% to 34%, $I^2=98$%), and other therapeutic management incidents (24%, 21% to 30%, $I^2=98$%), accounted for the highest proportion of preventable patient harm followed by incidents related to surgical procedures (23%, 9% to 38%, $I^2=98$%), healthcare infections (16%, 11% to 22%, $I^2=98$%), and diagnosis (16%, 11% to 21%, $I^2=98$%).

### Table 1 | Proportions of types of preventable patient harm and overall harm

| Outcome                  | Overall harm | Preventable harm |
|--------------------------|--------------|------------------|
| **Prevalence**           | % (95% CI)   | $I^2$            | Median (IQR) | % (95% CI) | $I^2$ | Median (IQR) |
| Overall                  | 6 (5 to 7)   | 99               | 5 (3-9)      | 12 (9 to 14) | 99    | 10 (7-15)    |
| Emergency department     | 6            | 3 (2 to 4)       | 78           | 3 (3-6)      | 5 (3 to 6) | 84 | 5 (4-6)      |
| Hospitals                | 45           | 5 (4 to 6)       | 99           | 5 (3-7)      | 10 (9 to 12) | 99 | 10 (7-12)    |
| Intensive care           | 6            | 18 (12 to 26)    | 96           | 14 (10-28)   | 34 (19 to 50) | 99 | 29 (20-59)  |
| Obstetrics               | 4            | 2 (0 to 4)       | 95           | NA           | 4 (2 to 6) | 92 | NA           |
| Primary care             | 3            | 3 (0 to 9)       | 93           | 7 (3 to 10)  | 0       | NA           |
| Surgery                  | 6            | 10 (7 to 13)     | 97           | 9 (9-10)     | 20 (14 to 27) | 99 | 22 (15-30)  |
| **Severity of patient harm** |             |                  |              |              |              |              |
| Mild                     | 20           | 49 (43 to 56)    | 97           | 45 (40-55)   | 50 (41 to 59) | 98 | 49 (43-58)  |
| Moderate                 | 20           | 36 (31 to 42)    | 96           | 38 (30-50)   | 36 (28 to 44) | 98 | 36 (27-47)  |
| Severe                   | 20           | 12 (9 to 15)     | 94           | 10 (8-19)    | 12 (8 to 15) | 95 | 13 (6-17)   |
| **Types of patient harm** |             |                  |              |              |              |              |
| Drugs                    | 25           | 25 (16 to 34)    | 98           | 20 (9-35)    | 26 (19 to 34) | 99 | 21 (17-30)  |
| Other therapeutic        | 17           | 24 (21 to 30)    | 98           | 22 (16-30)   | 20 (9 to 31) | 98 | 21 (12-32)  |
| Procedure                | 20           | 23 (13 to 33)    | 98           | 18 (6-28)    | 24 (17 to 31) | 98 | 19 (14-32)  |
| Surgical procedure       | 18           | 23 (9 to 38)     | 98           | 21 (8-36)    | 31 (20 to 42) | 98 | 27 (16-41)  |
| Healthcare infections    | 14           | 16 (11 to 21)    | 98           | 12 (9-22)    | 9 (6 to 12) | 98 | 10 (6-11)   |

The proportions for types of preventable or overall harm do not add to 100% because each figure in the table is the pooled proportion which has been calculated by combining (after assigning appropriate weights) proportions extracted from several independent studies using meta-analysis. Moreover, not all studies reported all types of preventable or overall harm and therefore it is not appropriate to assume they add up to 100%. NA=not applicable.
Fig 2 | Forest plot of the pooled prevalence of preventable patient harm across medical care settings
specialties such as surgery and intensive care ($b=0.08$, 95% confidence interval 0.05 to 0.11), in studies with relatively small sample sizes ($b=0.03$, 0.01 to 0.06), and in studies on children and older adults ($b=0.03$, −0.01 to 0.05). These three variables (medical care setting, population group, and sample size) were therefore eligible for inclusion in the multivariable regression analysis. All the other variables (research design, assessment method of harm, assessment of preventability, risk of bias, and WHO region) were ineligible for inclusion in multivariable analyses because none of them influenced the prevalence of preventable patient harm in univariable analyses ($P>0.10$).

The overall multivariable model was statistically significant ($\chi^2 (4)=33.98$, $P<0.001$) and reduced the $I^2$ statistic from 79% to 31%. Only the medical care setting ($b=0.07$, 95% confidence interval 0.04 to 0.10) remained a significant predictor of the prevalence of preventable patient harm in multivariable analyses suggesting that the prevalence of preventable patient harm is higher in advanced medical specialties (surgery and primary care) compared with studies in general hospitals. The population group and sample size were not significantly associated with the prevalence of preventable patient harm after controlling for the medical care setting in the multivariable analyses.

Small study bias

Figure 3 shows some evidence of publication bias as indicated by visual inspections of the funnel plots and by the Egger test for small study effects for the primary outcome (bias coefficient for the main analysis 1.20, 95% confidence interval 0.24 to 2.15, $P=0.02$).

**Discussion**

Understanding and mitigating preventable patient harm is a major public health challenge across the globe. We conducted a systematic review and meta-analysis to understand the prevalence, severity, and common types of preventable patient harm across medical care settings. We pooled data from 70 studies and found that preventable patient harm occurs in 6% of patients across medical care settings. Considering that a pooled prevalence of 12% for overall harm was found, we conclude that half of patient harm is preventable. The proportion of severe preventable patient harm causing prolonged, permanent disability or death was 12%. The most common types of preventable patient harm were related to drugs, other therapeutic management, and invasive medical and surgical procedures. The most extensive evidence on preventable patient harm comes from hospitals (45 studies) but less evidence is available for specific medical specialties. Preventable patient harm was more prevalent in patients treated in surgical and intensive care units compared with patients treated within across general hospitals. None of the other method variations which we examined across the studies influenced the pooled prevalence of preventable patient harm (population group, research design, assessment method of harm, assessment of preventability, sample size, risk of bias, or WHO region).

### Table 2 | Univariable and multivariable predictors of the prevalence of preventable patient harm (n=70)

| Variable                                      | Univariable | Multivariable |
|-----------------------------------------------|-------------|---------------|
| WHO region:                                   |             |               |
| US                                            | 33          | 1             |
| Europe                                        | 27          | −0.01 (−0.03 to 0.01) 0.01 0.59 NA NA NA |
| Asia or other                                 | 10          | −0.01 (−0.02 to 0.04) 0.02 0.54 NA NA NA |
| Medical setting:                              |             |               |
| General hospitals and obstetrics              | 49          | 1             |
| Primary care and emergency department         | 9           | −0.02 (−0.05 to 0.01) 0.02 0.18 −0.03 (−0.06 to 0.01) 0.02 0.12 |
| Advanced hospital specialties                 | 12          | 0.08 (0.05 to 0.11) 0.02 <0.001 0.07 (0.04 to 0.10) 0.01 <0.001 |
| Research design:                              |             |               |
| Retrospective or cross sectional              | 50          | 0.01 (−0.01 to 0.04) 0.01 0.31 NA NA NA |
| Prospective                                   | 20          | 0.01 (−0.01 to 0.04) 0.01 0.31 NA NA NA |
| Sample size:                                  |             |               |
| Large (>1000)                                 | 43          | 1             |
| Small (<1000)                                 | 27          | 0.03 (0.01 to 0.06) 0.01 0.02 0.02 (−0.01 to 0.04) 0.01 0.12 |
| Population:                                   |             |               |
| Adults                                        | 59          | 1             |
| Children or older adults                      | 11          | 0.03 (−0.01 to 0.05) 0.02 0.09 0.02 (−0.01 to 0.05) 0.01 0.09 |
| Assessment method:                            |             |               |
| Medical record review                         | 53          | 1             |
| Surveys with patients and health providers    | 17          | −0.01 (−0.04 to 0.02) 0.01 0.58 NA NA NA |
| Preventability by consensus among reviewers:  |             |               |
| Standardised Likert scale                     | 43          | 1             |
| Implicit criteria                             | 27          | 0.01 (−0.01 to 0.04) 0.01 0.36 NA NA NA |
| Risk of bias:                                 |             |               |
| High (<7 score)                               | 41          | 1             |
| Low (<7 score)                                | 29          | −0.01 (−0.03 to 0.02) 0.01 0.89 NA NA NA |

SE=standard error; NA=not applicable.
Strengths and limitations of the study

Despite the unique focus on preventable patient harm and several method strengths, this review has also limitations. Firstly, the prevalence of preventable patient harm varied considerably across studies and this variation was only partly explained in meta regression analyses. Other relevant factors likely accounted for the unexplained heterogeneity. For example, variations in the timeframe used to detect harm might be important when interpreting the differences in the prevalence estimates, along with variations in the implementation of quality assurance programmes and the quality of the documentation used for detecting preventable patient harm. For example, quality assurance programmes have possibly been implemented in parallel with some of the reviewed studies which might account for some proportion of the heterogeneity that we observed in this meta-analysis.

Secondly, a critical eligibility criterion to ensure feasibility of this review was that data on preventable patient harm were available in the published reports of the studies. Studies which did not report data on preventable patient harm were excluded from the analyses. However, most studies focused primarily on overall patient harm, reported preventable patient harm as a secondary outcome, and only one third of the studies provided an analysis of severity and types of preventable patient harm.

Thirdly, preventability rankings are likely to evolve over time especially after new technological advancements in healthcare. Consequently, some patient harms which are now considered non-preventable might be preventable in the future. However, the studies we reviewed consistently found that about 50% of patient harm was preventable and we did not observe any different patterns over the past 19 years.

Fourthly, over half of the reviewed studies employed retrospective case record reviews to investigate the prevalence, nature, and severity of preventable patient harm. Although case record reviews are the most universally used method for assessing patient harm to date, patients and healthcare providers have repeatedly expressed concerns that data contained in case records do not capture the full range of harms that they experience during their healthcare encounters. On the other hand, self reporting of patient harms (either by patients or healthcare providers) relies on recall and has its own limitations. Combining methods (such as prospective case record reviews with surveys with patients and healthcare providers) with the parallel engagement of patients as partners in identifying medical errors and mitigating preventable patient harm are promising approaches for enhancing patient safety.

Comparison with other studies

Our headline finding is that preventable patient harm is a highly prevalent international healthcare challenge which causes unnecessary patient suffering and can result in several avoidable deaths. As this review is specifically designed to understand patterns of preventable patient harm, comparisons with existing reviews focused on overall harm is problematic. Although we concur that examining the nature of overall harm is important, increasing the emphasis on preventable patient harm (which is the most amenable form of patient harm) is critical in terms of designing efficient patient safety strategies.

There is also evidence that preventable patient harm is not only a public health concern but incurs a considerable opportunity cost. The excess length of hospital stays attributable to medical errors is estimated to be 2.4 million hospital days, which accounts for $9.3 billion (£7.3bn; €8.2bn) excess charges in the US. Similarly, only six selected types of preventable patient harms in English hospitals result in 934 excell bed days per 100 000 population, which is equivalent to over 3500 salaried hospital nurses each year. Thus, investments in developing and evaluating mitigation strategies for preventable patient harm are urgently needed and are strongly supported by our findings.

Policy implications

Our findings provide a useful agenda of priority areas for mitigating preventable patient harm. When exploring the nature of preventable patient harm, drug related and therapeutic incidents comprise the majority. This finding echoes recommendations from international patient safety policy initiatives in the past decade including the recent WHO’s third global patient safety challenge “medication without harm.” Thus, it would be logical to prioritise efforts on developing and testing evidence-based mitigation strategies for these specific types of preventable patient harm. As this study establishes the scale of preventable patient harm in medical care settings, the need to gain better insight about the systemic and cultural circumstances under which preventable patient harm occurs is highlighted as a priority area. Several studies have sought to explain patient harms by reference to their sociotechnical context. For example, Vincent and colleagues proposes that patient harm occur because of contributory factors (which include “active” and “latent” failures) in the healthcare system. These failures correspond to
characteristics of the system such as the tasks that are undertaken, the people, technology, and tools that are involved, and the organisational values and structures in which the system operates. The studies included in our review, however, did not provide much insight into the way in which such factors might have contributed to the instances of preventable patient harm identified. Retrospective examination of patient harm often does not capture the myriad ways in which contributory factors could combine to produce—or avert—a preventable incident of patient harm. Mixed method approaches, which connect the occurrence of patient harm to the presence of specific contributory factors and engage patients as partners in establishing these connections, have excellent prospects to achieve an in depth understanding of possible pathways to patient harm.

A thorough understanding of the nature of preventable patient harm and its determinants could offer useful, evidence-based directions for designing efficient mitigation strategies. A combination of individual-level measures (eg, educational interventions for practitioners), system-level measures (eg, human-centred design of healthcare tasks and work environments), and organisational-level measures (eg, introducing quality monitoring and improvement processes) are likely to be a promising strategy for mitigating preventable patient harm, but scalable evaluations of these interventions are needed to support wider implementation. Furthermore, the interventions depend on the presence of an organisational context that supports their implementation.

Another important finding is that preventable patient harm appears to be a serious concern in advanced medical specialties including intensive care and surgical units. Patients treated in these specialties were more likely to experience preventable patient harm compared with patients treated in general hospitals. Surgical harm is a sizeable part of the overall in-hospital harm, but our estimates are higher than anticipated. The underlying causes of these figures warrant further investigation because current safety standards could “be failing to rescue” many high risk patients treated in advanced specialties. Moreover, clinicians in these specialties are often exposed to work pressures and are expected to deliver life-changing decisions quickly which might negatively impact on their personal wellbeing, a well known risk factor for preventable medical incidents.

On the other hand, surgery and intensive care units deal with high risk patients to whom complex medical procedures are implemented. Patient harm therefore might be more detectable in these settings because of its immediate, serious, or cumulative impact on patients’ health or because better surveillance systems for detecting patient harm are implemented in these settings. Additionally, it is not always clear from the study designs that some proportion of the preventable patient harm has not occurred in the transition between general hospital care and advanced specialty care.

Another major contribution of our synthesis is that it highlights key gaps in the literature on preventable patient harm. Only two studies were based in primary care, where over 80% of healthcare service is delivered internationally, and no evidence was identified in psychiatry. Certain types of preventable harms which tend to occur in primary care and psychiatry might remain undetected or untargeted by quality and safety improvement programmes. For example, we found that diagnostic harm is a common preventable type of harm but our understanding of its nature needs to be improved. A likely explanation is that diagnostic harm is directly or indirectly linked with the provision of services in primary care where research on preventable patient harm is sparse. Obtaining more precise estimates of the types and sources of preventable diagnostic harm occurring in primary care or in transitions from primary care to hospital care could lay the foundation for implementing efficient interventions for diagnostic harm. Systemic interventions, enhanced patient involvement in decision making for diagnoses, use of electronic tools, and emotion-cognitive interventions for boosting practitioners’ confidence or certainty in making diagnoses are potentially fruitful intervention areas for reducing diagnostic harm but have not been systematically evaluated or implemented in practice.

Less than a handful of studies focused on children and older adults, groups increasingly viewed as vulnerable to low quality or unsafe care. Furthermore, only a fraction of the included studies were conducted in developing countries, as many studies from developing countries failed to provide data on preventability of harm which rendered them ineligible. Thus, despite the evidence showing that the prevalence of overall harm is higher in developing countries compared with developed countries, we did not find such difference for preventable patient harm.

Commissioning research to understand the prevalence, nature, and impact of preventable patient harm in primary care and psychiatry, among vulnerable patient groups (eg, young children, older adults, or marginalised groups of the society such prison healthcare) and in developing countries has the potential to advance policy guidance and practice for mitigating preventable patient harm.

Conclusion

Our findings affirm that preventable patient harm is a serious problem across medical care settings. Priority areas are the mitigation of major sources of preventable patient harm (such as drug incidents) and greater focus on advanced medical specialties. It is equally imperative to build evidence across specialties such as primary care and psychiatry, vulnerable patient groups, and developing countries. Improving the assessment and reporting standards of preventability in future studies is critical for reducing patient harm in medical care settings.

Contributors: The original idea for the research was developed by MP, DMA, RNK, DP, PB, AJA. MP conducted the analysis with input from KK, EK, DMA, RNK, DP, AA, PB, and AJA. MP and KK conducted the searches, study selection, quality assessments, and other data
Towards an international patient safety taxonomy. World Health Organization, 2003.

14 Brennan TA, Leape LL, Laird NM, et al, Harvard Medical Practice Study I. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. 1991. Qual Saf Health Care 2004;13:145-51. discussion 151-2. doi:10.1136/qqhcs.2002.003822

15 Anderson O, Davis R, Hanna GB, Vincent CA. Surgical adverse events: a systematic review. Am J Surg 2013;206:253-62. doi:10.1016/j.ajss.2012.11.009

16 Long SJ, Brown KF, Ames D, Vincent C. What is known about adverse events in older medical hospital inpatients? A systematic review of the literature. Qual Saf Health Care 2013;22:542-54. doi:10.1093/qshc/mid056

17 Nilsson L, Phil A, Tägås M, Ericsson E. Adverse events are common on the intensive care unit: results from a structured record review. Acta Anesthesiol Scand 2012;56:959-65. doi:1111/j.1399-6576.2012.02711.x

18 Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-Analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2064-8. doi:10.1001/jama.283.15.2008

19 Lilford R, Stirling S, Maillard N. Citation classics in patient safety research: an invitation to contribute to an online bibliography. Qual Saf Health Care 2005;14:386-92. doi:10.1136/qshc.2005.017178

20 Luchini C, Stubbs B, Solmi M, et al. Assessing the quality of studies in meta-analyses: Advantages and limitations of the Newcastle Ottawa Scale. World J Med 2015;5:84-9. doi:10.1016/j.wjma.2014.08.001

21 Kontopantelis E, Reeves D. metaan: Random-effects meta-analysis. Stata J 2010;10:395-407. doi:10.1177/1536867X1001000307

22 Harbord RM, Higgins JP. Meta-regression in Stata. Stata J 2008;8:493-519. doi:10.1177/1536867X0800800403

23 Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1555-73. doi:10.1002/sim.1187

24 Freeman MF, Tukey JW. Transformations Related to the Angular and the Square Root. Ann Math Stat 1950;21:607-11. doi:10.1214/aoms/1177729576

25 DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88. doi:10.1016/0197-2456(86)90046-2

26 Ioannidis JP, Haidemenos G, Ntzani EEA, Trikalinos TA. Meta-analysis: problems and perspectives. Trends Pharmacol Sci 2008;29:83-7. doi:10.1016/j.tips.2007.12.001

27 Kontopantelis E, Reeves D. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: a comparison between DerSimonian-Laird and restricted maximum likelihood. Stat Methods Med Res 2012;21:657-9. doi:10.1177/0962280211413451

28 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60. doi:10.1136/bmj.327.7414.557

29 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:160-4. doi:10.1136/bmj.315.7109.629

30 Hedges LW, Olkin I. Statistical Methods for Meta-Analysis. Orlando, Florida: Academic Press, 1985.

31 Koo T, Li Y. A new way to get confidence intervals for Cohen’s d. J Educ Behav Stat 2012;37:255-67. doi:10.3203/jes-2012-0272

32 Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1555-73. doi:10.1002/sim.1187

33 Harbord RM, Higgins JP. Meta-regression in Stata. Stata J 2008;8:493-519. doi:10.1177/1536867X0800800403

34 Toh H, Turner B, Aiken LH, Sammel MD. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1555-73. doi:10.1002/sim.1187

35 Hedges L, Mirkovic DM, Olkin I. A new way to get confidence intervals for Cohen’s d. J Educ Behav Stat 2012;37:255-67. doi:10.3203/jes-2012-0272

36 DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88. doi:10.1016/0197-2456(86)90046-2

37 Ioannidis JP, Haidemenos G, Ntzani EEA, Trikalinos TA. Meta-analysis: problems and perspectives. Trends Pharmacol Sci 2008;29:83-7. doi:10.1016/j.tips.2007.12.001

38 Kontopantelis E, Reeves D. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: a comparison between DerSimonian-Laird and restricted maximum likelihood. Stat Methods Med Res 2012;21:657-9. doi:10.1177/0962280211413451

39 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60. doi:10.1136/bmj.327.7414.557

40 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:160-4. doi:10.1136/bmj.315.7109.629

41 Hedges LW, Olkin I. Statistical Methods for Meta-Analysis. Orlando, Florida: Academic Press, 1985.

42 Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1555-73. doi:10.1002/sim.1187

43 Harbord RM, Higgins JP. Meta-regression in Stata. Stata J 2008;8:493-519. doi:10.1177/1536867X0800800403

44 Toh H, Turner B, Aiken LH, Sammel MD. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1555-73. doi:10.1002/sim.1187

45 Harbord RM, Higgins JP. Meta-regression in Stata. Stata J 2008;8:493-519. doi:10.1177/1536867X0800800403
39 Baines RJ, Langelaa M, de Brujinie MC, et al. Changes in adverse event rates in hospitals over time: a longitudinal retrospective patient record review study. BMJ Qual Saf 2011;22:290-8. doi: 10.1136/bmjqs-2012-001126

40 Baker GR, Nentwich MA, Landgraf J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ 2004;170:1678-86. doi: 10.1503/cmaj.1040498

41 Bartlett G, Blais R, Tamblyn R, Clermont R, MacGibbon B. Impact of patient communication problems on the risk of preventable adverse events in acute care settings. CMAJ 2008;178:1555-62. doi: 10.1503/cmaj.070690

42 Blais R, Sears NA, Doran D, et al. Assessing adverse events among home care clients in three Canadian provinces using chart review. BMJ Qual Saf 2013;22:989-97. doi: 10.1136/bmjqs-2013-002039

43 Calder L, Pozgay A, Riff S, et al. Adverse events in patients with return emergency department visits. BMJ Qual Saf 2015;24:142-8. doi: 10.1136/bmjqs-2014-003194

44 Calder LA, Forster A, Nelson M, et al. Adverse events among patients registered in high-acuity areas of the emergency department: a prospective cohort study. CJEM 2010;12:421-30. doi: 10.1017/...
retrospective case-note review. Age Ageing 2008;37:265-9. doi:10.1093/ageing/afo43

86 Sani AB-A, 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. doi:10.1136/qhc.2006.021154

87 Sop M, Fryksmark U, Köster M, Haglund B. The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. Int J Qual Health Care 2009;21:285-91. doi:10.1093/intqhc/mzp025

88 Soura P, Uva AS, Sermanheira F, Nunes C, Leite ES. Estimating the incidence of adverse events in Portuguese hospitals: a comparison improving quality and patient safety. BMC Health Serv Res 2014;14:311. doi:10.1186/1475-9269-14-311

89 Stockwell DC, Bisanya H, Classen DC, et al. A trigger tool to detect harm in inpatient inpatient settings. PEDIATRICS 2015;135:1036-42. doi:10.1542/peds.2014-2152

90 Stockwell DC, Landrigan CP, Toomey SL, et al. GAPPs Study Group. Adverse Events in Hospitalized Pediatric. Pediatrics 2018;142.e20173360. doi:10.1542/peds.2017-3360

91 Suarez C, Menendez MD, Alonso J, Castaño N, Alonso M, Vazquez F. Detection of adverse events in an acute geriatric hospital over a 6-year period using the Global Trigger Tool. J Am Geriatr Soc 2014;62:896-900. doi:10.1111/jgs.12774

92 Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. BMJ 2000;320:741-4. doi:10.1136/bmj.320.7237.741

93 Vincent C, Neale G, Wolschonshwyk M. Adverse events in British hospitals: preliminary retrospective record review study. BMJ 2001;322:517-9. doi:10.1136/bmj.322.7285.517

94 Weingart SN, Pagovich O, Sands DS, et al. What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. J Gen Intern Med 2005;20:830-6. doi:10.1111/j.1525-1497.2005.0180.x

95 Williams DJ, Olsen S, Crichton W, et al. Detection of adverse events in a Scottish hospital using a consensus-based methodology. Scott Med J 2008;53:26-30. doi:10.1258/rsmh.53.4.26

96 Wilson RM, Michiel P, Olsen S, et al. WHO Patient Safety CRM/AFRO Working Group. Patient safety in developing countries: retrospective estimation of scale and nature of harm to patients in hospital. BMJ 2012;344:e875. doi:10.1136/bmj.e875

97 Woods D, Thomas E, Holl J, Altman S, Brennan T. Adverse events and preventable adverse events in children. Pediatrics 2005;115:155-60. doi:10.1542/peds.2004-0410

98 Zegers A, de Bruyne MC, Wagner C, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. Qual Saf Health Care 2009;18:297-302. doi:10.1136/qhc.2007.025924

99 Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and deaths in hospital. Results of the Harvard Medical Practice Study I. N Engl J Med 1991;324:370-6. doi:10.1056/NEJM199102073240604

100 Guarousse-Deroges M, Odsi, Melissari E, Filiapi E, Max A, Lau N, Misset B. Overview of medical errors and adverse events. Ann Intensive Care 2022;2:2. doi:10.1186/s13611-021-00577-z

101 Schwappach DLP. Review: engaging patients as vigilant partners in safety: a systematic review. Med Care Res Rev 2010;67:119-48. doi:10.1177/0744842810375524

102 Ricci-Cabelllo I, Avery AJ, Reeves D, Kadam UT, Valderas JM. Measuring and Excess Bed-Days Due to 6 Patient Safety Incidents: Empirical analysis of randomized controlled trials. BMJ Qual Saf 2012;21:160-70. doi:10.1136/bmjqs-2011-000150

103 Phillips DL, Walshie K, Parker D, Noyce PR, Ashcroft DM. Job characteristics, well-being and risky behaviour amongst pharmacists. Psychol Health Med 2016;21:92-4. doi:10.1080/13548506.2014.884001

104 Huiskes VJ, Burger DM, van den Dode EH, van den Bilt BJ. Effectiveness of medication review: a systematic review and meta-analysis of randomized controlled trials. BMJ Fam Pract 2017;18:5. doi:10.1136/bmjpsr-2017-000128

105 Khalil H, Bell B, Chambers H, et al. Professional, structural and organisational interventions in primary care for reducing medication errors. Cochrane Database Syst Rev 2017;10003942. doi:10.1002/14651858.CD003942.pub3

106 Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation’s programme evaluations and relevant literature. BMJ Qual Saf 2012;21:167-84. doi:10.1136/bmjqs.2011-000760

107 Phillips DL, Jones CE, Parker D, Ashcroft DM. Organizational conditions for engagement in quality and safety improvement: a longitudinal qualitative study of community pharmacies. BMC Health Serv Res 2012;12:303. doi:10.1186/1472-6963-12-303

108 Graham KE, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. N Engl J Med 2009;361:1368-75. doi:10.1056/NEJMoa0903048

109 Shanafelt TD, Balch CM, Beckgams G, et al. Burnout and medical errors among American surgeons. Ann Surg 2010,251:955-1000. doi:10.1097/SLA.0b013e3181fbbdab

110 Panesar SS, De Silva D, Carson-Stevens A, et al. How safe is primary care? A systematic review. BMJ Qual Saf 2016;25:544-53. doi:10.1136/bmjqs-2015-004478

111 Zwaan L, Schiff GD, Singh H. Advancing the research agenda for diagnostic error reduction. BMJ Qual Saf 2013;22 Suppl 2:i52-7. doi:10.1136/bmjqs.2012-001624

112 Riches N, Panagioti M, Alam R, et al. The Effectiveness of Electronic Differential Diagnoses (DDX) Generators: A Systematic Review and Meta-Analysis. PloS One 2016;11:1.e0148991. doi:10.1371/journal.pone.0148991

113 Singh H, Graber ML, Kissen SM, et al. System-related interventions to reduce diagnostic errors: a narrative review. BMJ Qual Saf 2012;21:160-70. doi:10.1136/bmjqs-2011-000500

114 McIlroy H, van der Doelen-Louwianidts D, et al. Patient safety strategies targeted at diagnostic errors: a systematic review. Ann Intern Med 2013;158:381-9. doi:10.7326/0003-4819-158-5-20130501-00004

Supplementary materials: Searches and eTable 1-4