OPEN ACCESS

Incidence and characteristics of adverse events in paediatric inpatient care: a systematic review and meta-analysis

Pernilla Dillner ,[1,2] Luisa C Eggenschwiler ,[3] Anne W S Rutjes ,[4,5] Lena Berg ,[6,7] Sarah N Musy ,[3] Michael Simon ,[3] Giusi Moffa ,[8] Ulrika Förberg ,[1,6] Maria Unbeck [6,9]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjqs-2022-015298).

For numbered affiliations see end of article.

Correspondence to
Pernilla Dillner, Department of Women's and Children's Health, Karolinska Institutet, Stockholm 171 77, Sweden;
pernilla.dillner@ki.se

PD and LCE are joint first authors.

Received 23 June 2022
Accepted 8 November 2022

To cite: Dillner P, Eggenschwiler LC, Rutjes AWS, *et al*. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2022-015298

ABSTRACT

Background Adverse events (AEs) cause suffering for hospitalised children, a fragile patient group where the delivery of adequate timely care is of great importance.

Objective To report the incidence and characteristics of AEs, in paediatric inpatient care, as detected with the Global Trigger Tool (GTT), the Trigger Tool (TT) or the Harvard Medical Practice Study (HMPS) method.

Method MEDLINE, Embase, Web of Science and Google Scholar were searched from inception to June 2021, without language restrictions. Studies using manual record review were included if paediatric data were reported separately. We excluded studies reporting: AEs for a specific disease/diagnosis/treatment/procedure, or deceased patients; study protocols with no AE outcomes; conference abstracts, editorials and systematic reviews; clinical incident reports as the primary data source; and studies focusing on specific AEs only. Methodological risk of bias was assessed using a tool based on the Quality Assessment Tool for Diagnostic Accuracy Studies 2. Primary outcome was the percentage of admissions with ≥1 AEs. All statistical analyses were stratified by record review methodology (GTT/TT or HMPS) and by type of population. Meta-analyses, applying random-effects models, were carried out. The variability of the pooled estimates was characterised by 95% prediction intervals (PIs).

Results We included 32 studies from 44 publications, conducted in 15 countries totalling 33 873 paediatric admissions. The total number of AEs identified was 8577. The most common types of AEs were nosocomial infections (range, 6.8%–59.6%) for the general care population and pulmonary-related (10.5%–36.7%) for intensive care. The reported incidence rates were highly heterogeneous. The PIs for the primary outcome were 3.8%–53.8% and 6.9%–91.6% for GTT/TT studies (general and intensive care population). The equivalent PI was 0.3%–33.7% for HMPS studies (general care). The PIs for preventable AEs were 7.4%–96.2% and 4.5%–98.9% for GTT/TT studies (general and intensive care population) and 10.4%–91.8% for HMPS studies (general care). The quality assessment indicated several methodological concerns regarding the included studies.

Conclusion The reported incidence of AEs is highly variable in paediatric inpatient care research, and it is not possible to estimate a reliable single rate. Poor reporting standards and methodological differences hinder the comparison of study results.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The only available systematic review in this area is dated and shows a surprisingly low estimate of adverse event (AE) incidence. As paediatric inpatients are particularly vulnerable and run a high risk of exposure to AEs, a systematic review examining this important knowledge gap is lacking.

WHAT THIS STUDY ADDS

⇒ This review gives an up-to-date estimate of the incidence and variation of paediatric inpatient AEs. It also adds relevant methodological reflections about structured retrospective record review methods, as well as their application and reporting quality.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ A better knowledge of the complex nature of paediatric AEs is important for the development of more targeted patient safety interventions to increase quality of care and prevent paediatric patients suffering AEs. An awareness of the current incomplete reporting of key elements related to AE data may help researchers to improve the quality of reporting in future studies.

INTRODUCTION

Adverse events (AEs) are costly,[1] cause suffering for patients, their families and for healthcare professionals[2] and have been recognised as a critical global healthcare issue.[3 4] An AE may be defined as unintended physical injury

resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalisation, or that results in death.[5] The incidence of AEs varies between contexts (eg, country, hospital types, included specialities) and research is heavily influenced by the method used. Between 7% and 40% adult general inpatients are affected by AEs. These are often deemed to be preventable, indicating that patient safety can be improved.[6]

Hospitalised children are a fragile patient group. Even a low degree of error related to acts of omissions or commissions can affect the child's health and in the long-term risk affecting the child's development and future.[7] Patients treated at intensive care units run a greater risk of being exposed to AEs than general care patients.[8 9] Sedation and the need for intravascular and/or breathing devices are factors associated with AEs in paediatric patients. Those patients experiencing AEs are on average younger and have a longer length of stay.[8]

There are various methods for detecting, measuring and characterising AEs in healthcare, but as yet no gold standard exists.[10] A commonly used method is structured retrospective record review, which includes different approaches, for example, the Harvard Medical Practice Study (HMPS) method[11 12] or the Global Trigger Tool (GTT)[5] with its subsequent adaptations (Trigger Tools, TTs) to be used in different contexts, such as paediatrics,[13 14] oncology,[15] psychiatry[16] or home healthcare.[17] Record review has been shown to be superior in detecting AEs compared with other methods, such as incident reporting systems and administrative data.[14 18–20]

In adult care, several systematic reviews[6 21–24] regarding the identification of AEs using record review methodology, with or without meta-analysis, have been published. To the best of our knowledge, only one systematic review focusing on paediatric care has been published.[25] This review included nine publications, of which six used record review data and three used administrative record data, and was restricted to a minimum of 1000 patients. The admission year for included patients ranged from 1984 to 2009. This review presents a surprisingly low AE incidence. The publications of GTT and TT studies in the paediatric context have increased in the last 10 years. Therefore, an updated systematic review, irrespective of study sample sizes, was indicated. The aim of this systematic review is to report the incidence and characteristics of AEs, in paediatric inpatient care, as detected with the GTT, the TT or the HMPS method.

METHODS

The review was carried out as a systematic review and meta-analysis. The study protocol was uploaded on h 10.5281/zenodo.5513354.

Information sources and search strategies

The following databases were used for the search: MEDLINE, Embase, Web of Science and Google Scholar.[26] A search strategy was developed with the help of librarians, and this encompassed subject headings and free text words that described the population, the context, the concept and type of evidence source. The search terms used were: *Iatrogenic Disease, Medical Errors, Patient Harm, adverse event*, harm, trigger*, Adolescent, Child, Infant, p?ediatric*, neonat*, child*, newborn* infant*, adolescen*, premature*, preschool, teenager*, Hospitals, Inpatients, Hospitalisation, Hospital Units, Hospital Departments, hospital*, intensive care, inpatient*, review*, record*, chart*, trigger tool* and *Harvard Medical Practice**. The systematic searches were performed between 4 and 8 June 2021 and no restrictions in language or publication year were applied. The full search strategy and outcomes for the respective database are shown in online supplemental material 1, tables S1–S4. Furthermore, the search was supplemented in the data extraction process with a manual scan of the reference lists of eligible publications.

Selection process

Publications that met the following criteria were included: (1) Children, all age groups, if cared for in paediatric inpatient units; (2) Studies including both adults and paediatric patients if the data for paediatric patients were reported separately; (3) Peer reviewed full text primary publications, reporting relevant quantitative outcome data; (4) Studies applying manual retrospective medical record review using GTT, TT or HMPS methodologies. We accepted all types of AE definitions (online supplemental material 1, table S5).

The following exclusion criteria were applied: (1) Publications reporting AEs for paediatric patients with a specific disease/diagnosis/treatment/procedure or who were deceased; (2) Studies in primary care, psychiatric care, day care/ambulatory care and emergency departments or other outpatient units at the hospital; (3) Study protocols without AE outcome data; (4) Publications such as conference abstracts, editorials and systematic reviews; (5) Studies using, for example, clinical incident reporting systems as the primary data source where these incident reports were subsequently analysed using record review; and (6) Publications reporting only specific AEs, for example, adverse drug events (online supplemental material 1, table S5).

The first screening step of applying the eligibility criteria to titles and abstracts was done independently by four reviewers, working in pairs (MU/PD, UF/LB). Thereafter, eligible full texts were retrieved, and the same reviewers independently assessed full texts. The reason for exclusion was noted and any discrepancies between the individual reviewers were discussed in the pairs until consensus was reached. If

required, discussion was held with the whole research group. Discussions during the selection process mostly concerned whether multiple publications on the same study were considered as overlapping or not.

Data extraction process

To ensure quality, data were independently extracted by two researchers per publication. Data regarding key study characteristics (eg, sample size, setting, number of hospitals, method used, patient demographics) and patient outcomes (incidence, frequencies, preventability, types, severity) were collected. Authors of 27 primary studies were contacted by email to request additional information to calculate the primary outcome or part of the secondary outcomes. Information was provided from 17 studies. Any discrepancies between reviewers were resolved in the same way as in the selection process and a consensus for each study was reached. All the studies included were discussed at some point within the research group. Discussions were either related to the quality assessment, the methodology or interpretation of data.

Quality assessment

To assess the methodological quality of each included study, a previously used quality assessment tool (QAT) was adapted. This QAT was based on the structure of the Quality Assessment Tool for Diagnostic Accuracy Studies 2 tool[27] and the content of the QAT by Musy *et al*[28] and later by Eggenschwiler *et al*[29] (online supplemental material 1, page 14). The QAT consists of five domains: patient selection, reviewers, record review process, outcomes and flow. Each domain includes two to three signalling questions which form the basis for the assessment of risk of bias and applicability-related concerns. These were rated as either low, high or unclear. Expert knowledge in quality assessments and record review methodology guided the adaptations. Examples of adaptations used were revisions of the domain record review process with signalling questions regarding support and monitoring during the review process. Furthermore, the risk of bias and applicability-related concerns were also rated as an overall judgement for each study (online supplemental material 2). The QAT for each study was used by two reviewers independently and a consensus was reached.

Primary outcome

A meta-analysis was carried out with the percentage of admissions with ≥1 AEs as the primary outcome measure.

Secondary outcomes

Secondary outcomes were AEs per 100 admissions, AEs per 1000 inpatient days, percentage of preventable AEs, as well as percentage of admissions with preventable AEs. In addition, types of AEs and AE severity were described.

Statistical analysis

Analyses were conducted using R V.4.1.3 on Linux[30] with the meta[31] and metafor[32] packages. All statistical analyses were stratified, distinguishing general and intensive care populations, as these are known from the literature to differ in the distribution of AEs.[6 22 33] They were also stratified by the record review methodology used (GTT/TT or HMPS). The categorisation of the two populations was based on whether most patients were admitted to either general or intensive care units. Studies using the HMPS methodology did not predominantly include intensive care patients. The GTT and TT methodologies were analysed together, as these methods share the same conceptual approach.

Where not explicitly reported, we calculated the number of admissions with ≥1 AE from the reported percentage estimates of admissions with AEs. Similarly, we derived the number of patient days by dividing the total number of AEs by the reported rate of AEs per 1000 patient days. Studies using the HMPS methodology were excluded from the meta-analyses for AEs per 100 admissions and AEs per 1000 patient days. Most of these studies included only the most severe AE per admission and therefore the estimates were not comparable.

We fitted random intercept logistic models, using the R metaprop function with the Wilson method for CIs for the meta-analysis of the percentage of admissions with ≥1 AE, the percentage of preventable AEs and the percentage of admissions with preventable AEs.[31] For the AEs per 100 admissions and AEs per 1000 patient days we used random intercept Poisson models, fitted with the R metarate function.[32]

Other systematic reviews on the same topic reported I^2 values of up to 100%[20 21 23]. Although frequently reported I^2 is not valid in the context of single proportions. We decided to characterise the variability of the estimates by reporting prediction intervals (PIs).[34 35] The 95% PI quantifies the sample variability and is expected to capture estimates from future studies with a 95% level of confidence.[36] We identified high heterogeneity, illustrated by the width of the PIs, which is wider than the 95% CI in the presence of between-study heterogeneity. Hence, we focused our reporting on PIs rather than CIs. Furthermore, we investigated heterogeneity via stratified analyses of five elements relating to risk of bias and four connected to applicability-related concerns. P values, derived from the likelihood ratio test for model fit, were considered statistically significant with a value of $p<0.05$. The PRISMA 2020 guideline for reporting systematic reviews was applied.[37]

RESULTS

Publication retrieval

The database searches yielded 3790 publications of which 1317 were duplicates leaving 2473 unique publications which were screened by title and

abstract. In total, 108 publications underwent full text screening, including four publications from reference lists. After assessment of eligibility, 64 publications were excluded and 44 publications[8 11–14 33 38–75] of 32 unique studies[8 11 12 14 33 38–42 45 48–50 53 55 57–68 71–74] were included (online supplemental material 1, figure S1). As one study[55] reported outcomes for both populations, a total of 33 samples were included, 22 for the general care and 11 for intensive care populations.

Study characteristics

The studies were published between 1991 and 2021 with inclusion periods ranging from year 1984 to 2019 and 59.4% of the studies were published in the last 10 years. The study periods ranged from 1 month to 6 years. The 32 studies originated from 15 countries, of which 34.4% were from North America, 28.1% from Europe, 18.8% from South America, 9.4% from Australia, 6.3% from Africa and 3.1% from Asia. In total, 33 873 paediatric admissions (median, 330; range: 11–6661) and 124 800 patient days (median, 2743; range, 87–21 789) were included. A wide variation of units was found, and 68.8% (n=22) of the 33 samples included mainly general care (eg, surgical, medical) and 34.4% (n=11) included mainly intensive care units for paediatric and neonatal patients. Patients' mean age (n=14 studies) varied between 3.0 years to 7.8 years and mean length of stay (n=17 studies) 2.8 days to 22.8 days (table 1). Most of the studies (n=28) were written in English, three in Spanish and one in Portuguese.

Study methodology characteristics

A majority used GTT/TT (n=23, 71.9 %), followed by HMPS (n=9, 28.1%). No study published after 2014 used the HPMS method. The most frequent sampling strategy was random (n=26, 81.3%). A majority of the 32 studies (n=25, 78.1%) were assessed to have used a two-stage retrospective record review process and the number of triggers/screening criteria varied between 14 and 88. Twenty-six (81.3%) described training prior to the review process (table 2) and 12 studies used test records.

Seven studies (21.9%) had teams where the whole or part of the team had prior experience in record review methodology and seven studies (21.9%) reported support during the review process, such as expert consultation. Ten studies (31.3%) described a monitoring process to ensure completeness, consistency and accuracy (data not shown).

Both acts of omissions and commissions were included in 53.1% (n=17) of the studies and 78.1% (n=25) included ≥1 AE per patient. Outcomes for inter-rater reliability, using double reviews, were reported in 53.1% (n=17). Of those, kappa values were reported in five (26.3%) studies, percentage agreement in four (21.1%), and both measures in eight (42.1%). Half of the studies included AE(s) that occurred both before, during and after index admission, and eight studies (25.0%) didn't specify the time frame for inclusion. The GTT manual's AE definition or similar was used in 17 studies (53.1%) and the HMPS definition or similar in 10 (31.3%) (table 2), and 77.8% (n=25) of these had a reference to their AE definition. Preventability was assessed in 19 studies (59.4%) (table 2).

AE descriptions

The total number of identified AEs was 8577 (range 0–34 per patient) in 33 samples, 3459 (range 0–27 per patient) in the general care population (13 GTT/TT and 9 HMPS samples) and 5118 (range 0–34 per patient) in the intensive care population (11 samples). Preventability was reported in 16 samples (48.5%) with a total of 3785 identified preventable AEs (online supplemental material 1, table S6).

The most common types of AEs in general care (n=9 studies) were nosocomial infections (range, 6.8%–59.6%), medication-related (2.3%–48.6%) and surgical-related (0.9%–30.5%). Pulmonary-related (10.5%–36.7%), nosocomial infections (6.6%–40%) and medical technical product-related (1.3%–30.8%) were the most common types of AEs in intensive care (n=8 studies) (table 3).

Twenty-one studies assessed and described the severity of paediatric AEs. A majority of these (71.4%, n=15) used a modified version of the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Scale (online supplemental material 1, table S6). The studies assessing severity by the modified NCC MERP Scale, irrespective of population, had a range for minor consequences, category E, between 16.5% and 88.4% (mean, 56.9%); major, category F, 0.0% and 62.7% (28.9%); permanent, category G, 0.0% and 14.8% (4.0%); life-threatening, category H, 0.0% and 28.9% (7.4%) and death, category I, 0.0% and 15.7% (2.7%). The intensive care population had a mean of 11.7% for the two most severe categories—life-threatening and death, whereas the general care population had a mean of 3.1%.

Meta-analyses

The forest plot in figure 1 shows the primary outcome, that is, percentage of admissions with ≥1 AEs for 32 out of 33 samples. The range of percentage of admissions ≥1 AEs for GTT/TT was 6.1%–38.0% and 16.2%–83.9% for general and intensive care and the equivalent for HMPS was 0.0%–19.0%. The pooled estimates for the GTT/TT (general and intensive care populations) were 17.7% (95% PI 3.8%–53.8%) and 47.3% (95% PI 6.9%–91.6%), respectively, and 3.9% (95% PI 0.3%–33.7%) for the HMPS (general care). There was a statistically significant difference in the pooled estimates between the two populations within the GTT/TT methodology (p=0.0003).

Table 1 Study characteristics

First author, publication year, country	Hospitals	Paediatric inpatients admissions, n	Inclusion period, year (months)	Type of hospital(s)	Academic level of hospital(s)	Inclusion and exclusion criteria of patients	Type of included units	Patient days, n	Age, (years) mean	LOS, (days) mean
General care population, GTT/TT methodology										
Chapman 2014[40] GBR	Multicentre	3992	2008–2011 (46)	Mixed	Mixed	LOS >24 hours	Mixed—not explicitly stated	NS	NS	NS
Davenport 2017[41] ARG	Single centre	200	2013 (12)	Paediatric	NS	LOS ≥48 hours, if >1 hospitalisation the most recent one was included, no psychiatric patients and not for social reasons	ICU, neonatology, general care (multipurpose unit)	1690	4.4	8.5
Fajreldines 2019[45] ARG	Single centre	318	2015–2016 (13)	Tertiary care	Academic	LOS ≥48 hours, patient s<18 years	Neonatal care, PICU, nursery, paediatric	2257	3.0	NS
Kirkendall 2012[50] USA	Single centre	240	2009 (12)	Paediatric	Academic	LOS ≥24 hours, any age, no psychiatric and rehabilitation patients	Mixed—not explicitly stated	1206	7.8	5.1
Matlow 2012[54–56] CAN	Multicentre*	3552	2008–2009 (12)	Mixed*	Mixed*	LOS ≥24 hours, patients <19 years, no obstetrics, or psychiatric patients and external transfers (except newborns)	Surgical, internal medicine, emergency, maternal/obstetrics, other	14738†	4.1†*	4.5†*
Paredes Esteban 2015,[58] ESP	Single centre	95	2014 (12)	NS	NS	Patients admitted to paediatric surgery, no patients with adverse events as the reason for admission	Surgical	406	6.7	4.2
Salimath 2020[46 60] IND	Single centre	520	NS (26‡)	Acute care	Academic	LOS≥24 hours, patients≤18 years, no psychiatric and rehabilitation patients	NICU, PICU, medicine, surgical, emergency and trauma†	2743†	NS	NS
Shah 2009[51 61] USA	Single centre	50	NS	Paediatric	Academic	Patients admitted to the otolaryngology service	Otolaryngology	87	NS	NS
Solevag, 2014[63] NOR	Single centre	494‡	2011 (3)	Acute care	Academic	Patients <18 years	Orthopaedic, surgical, ear/nose/throat, medicine	2001†	6.8†	4.1†
Stockwell 2015[66] USA	Multicentre	600	2012 (1)	Paediatric	Academic	LOS between 24 hours and 6 months, patients <22 years, no rehabilitation, normal newborn nursery, day treatment, psychiatric or obstetric patients	NS	4372	6.2	7.3
Stockwell 2018[13 67] USA	Multicentre	3790	2007–2012 (72)	Mixed	Mixed	LOS ≥24 hours, patients age <18 years, no psychiatric (without a concurrent acute medical issue) or rehabilitation patients	Mixed—not explicitly stated	21 789	NS	3.0§
Stroupe 2018[68] USA	Single centre	100	2014 (12)	Paediatric	Academic	Admitted patients ≤18 years	General paediatric, surgery, PICU, other	411	6.7	3.8
Unbeck 2014[14] SWE	Single centre	600	2010 (12)	Acute care	Academic	LOS ≥24 hours, patient <19 years	Neonatal, surgical/orthopaedic, medicine, emergency medicine	5559	4.3	9.3
General care population, HMPS methodology										
Brennan 1991[11 43 47 48 52] USA	Multicentre¶	6661**	1984 (12)	Acute care	Mixed¶	Admitted patients, no psychiatric patients	All types¶	NS	NS	NS
Davis 2002[42–44] NZL	Multicentre¶	1349†**	1998 (12)	Acute care	Mixed¶	Admitted patients, no day, psychiatric and rehabilitation-only patients	Mixed—not explicitly stated¶	4134‡	3.1†	3.1†
Letaief 2010[53] TUN	Single centre	116‡**	2005 (12)	Public	Academic	Admitted patients	Mixed—not explicitly stated	NS	NS	NS
Requena 2011[59] ESP	Multicentre	665	NS (NS)	NS	Mixed	LOS >24 hours, had a clinical history in the selected hospitals	NS	3318	NS	NS
Sommella 2014[64] ITA	Single centre	11‡**	2008 (12)	Acute care	NS	LOS >24 hours, no day hospital discharges	Medical, surgical, ICU¶	NS	NS	NS

Continued

Table 1 Continued

First author, publication year, country	Hospitals	Paediatric inpatients admissions, n	Inclusion period, year (months)	Type of hospital(s)	Academic level of hospital(s)	Inclusion and exclusion criteria of patients	Type of included units	Patient days, n	Age, (years) mean	LOS, (days) mean
Soop 2009[65] SWE	Multicentre ¶	159**	2003–2004 (12)	Acute care	Mixed¶	Admitted patients, no psychiatric, palliative care, rehabilitation, and day-only patients	Mixed—not explicitly stated¶	NS	NS	NS
Wilson 1995[12] AUS	Multicentre¶	2020**	1992 (12)	Acute care	Mixed¶	Admitted patients, no psychiatric and day-only patients	Different kind of medical and surgical†	8697†	4.1†	4.3†
Woods 2005[69 70 74] USA	Multicentre¶	3719**	1992 (12)	Profit/non-profit, government	Mixed¶	Admitted patients, no psychiatric, rehabilitation and drug/alcohol treatment patients	Mixed—not explicitly stated	NS	NS	NS
Zegers, 2009[33 75] NLD	Multicentre¶	330**	2004 (12)	Acute care	Mixed¶	LOS of >24 hours, no psychiatric, obstetrics and <1 year patients	Mixed—explicitly stated¶	NS	NS	NS
Intensive care population, GTT/TT methodology††										
Agarwal 2010[38] USA	Multicentre	734	2005 (4)	Mixed	Mixed	LOS ≥48 hours, no postoperative cardiac patients	PICU	5201	6.3	7.1
Barrionuevo 2010[39] ARG	Single centre	484	2006 (12)	Public	NS	LOS >24 hours	NICU	6465†	NS	13.4†
Hooper 2014[48] AUS	Single centre	59	2011 (3)	Paediatric	Academic	Admitted patients	PICU	164	NS	2.8
Jorro-Baron 2021[49] ARG	Multicentre	1465	2018–2019 (11)	Public	NS	LOS ≥24 hours, patients <18 years and admitted for acute care	PICU	15 842	4.6†	10.8†
Larsen 2007[8] USA	Single centre	259	2002–2003 (12)	Paediatric	Academic	Admitted patients	PICU	962†	NS	1.6§
Matlow 2012[54–56] CAN	Multicentre*	117	2008–2009 (12)	Mixed*	Mixed*	LOS ≥24 hours, patients <19 years, no obstetrics or psychiatric patients and external transfers (except newborns)	NICU, PICU	1574†	0.0†	13.5†
Maziero 2020[57] BRA	Multicentre	79	2017–2018 (NS)	Public	NS	Admitted patients	NICU, PICU	NS	NS	NS
Sharek 2006[62] USA/CAN	Multicentre	749	2004–2005 (3)	Mixed	Mixed	LOS ≥48 hours	NICU	17 106	NS	22.8
Ventura 2012[71] BRA	Single centre	218	2009 (6)	NS	NS	LOS ≥48 hours	NICU	2958	NS	13.5
Verlaat, 2018[72] NLD	Multicentre	48	2006–2012 (72)	NS	NS	LOS ≥2 hours, patient <18 years, no patients with corrected age <36 weeks (GA)	PICU	608†	6.4	12.7†
Vermeulen 2014[73] ZAF	Single centre	80‡‡	2012 (4)	Paediatric	Academic	LOS >48 hours, patients included only once if >1 admission	PICU	512	NS	4.0§

*Outcome for the total cohort.
†Additional data from authors.
‡Calculations are made.
§Median.
¶Information for the total cohort in a study with both paediatric and adult patients, information for the paediatric cohort not reported.
**Paediatric cohort.
††Studies using the HMPS methodology did not predominantly include intensive care patients.
‡‡Retrospective cohort.
Academic, academic medical centre/university hospital; Adm, admission; GA, gestational age; GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; ICU, Intensive care unit; LOS, length of stay; Mixed, both paediatric and adult hospital type; NICU, neonatal intensive care unit; NS, not specified; PICU, paediatric intensive care unit; TT, Trigger Tool.

Table 2 Study methodology

First author, publication year	RRR method	Method of record selection	Method of review*	> 1 AE/ patient	Commission/ omission	Data on IRR outcome	Time frame of AE detection, in addition to index admission	AE definition†	Preventability reporting
General care population, GTT/TT methodology									
Chapman, 2014[40]	TT	Random	2-stage; (1) 40 triggers, trained reviewer; (2) trained physician	Yes	Both	No	Before; after	Wider than IHI	No
Davenport, 2017[41]	GTT	Random	2-stage; (1) 52 triggers, trained physicians; (2) trained physicians	Yes	NS	Yes	Before NS; after	IHI like	No
Fajreldines, 2019[45]	TT	Random	NS	Yes	NS	No	NS	NS	No
Kirkendall, 2012[50]	GTT	Random	2-stage; 1) 53 triggers, trained nurses; (2) physician	Yes	Commission only	Yes	Before; after	IHI like	No
Matlow, 2012[54–56]‡	TT	Random	2-stage; (1) 35 triggers, trained nurses, health record technologists/medical record technicians; (2) trained physicians	Yes	Both	Yes	Before; after	HMPS like	Yes
Paredes Esteban, 2015[58]	GTT	Unclear	2-stage; (1) NS, trained nurses; (2) trained physicians	Yes	NS	No	Only index	HMPS like	No
Salimath, 2020[46 60]	TT	Random	2-stage; (1) 40 triggers, trained pharmacists; (2) trained physicians	Yes	NS	No	Before NS; after	Other	No
Shah, 2009[51 61]	TT	Random	2-stage; (1) 43 triggers, physicians; (2) physicians	Yes	Both	Yes	Before; after NS	Wider than IHI	No
Solevag, 2014[63]	TT	Convenience	Unclear; 39 triggers, trained physician	Yes	NS	No	Before; after	Wider than IHI	No
Stockwell, 2015[66]	TT	Random	2-stage; (1) 51 triggers, trained nurses, pharmacists; (2) trained physicians	Yes	Both	No	Before NS; after	IHI like	Yes
Stockwell, 2018[13 67]	TT	Random	2-stage; (1) 27 triggers, trained nurses; (2) trained physicians	Yes	NS	Yes	Before; after	IHI like	Yes
Stroupe, 2018[68]	TT	Random	2-stage; (1) 54 triggers, trained nurses; (2) trained physicians	Yes	Both	Yes	Before NS; after	Wider than IHI	Yes
Unbeck, 2014[14]	TT	Random	2-stage; (1) 88 triggers, trained nurses; (2) trained physicians	Yes	Both	Yes	Before; after	Wider than IHI	Yes
General care population, HMPS methodology									
Brennan, 1991[11 47 52]	HMPS	Random	2-stage; (1) 18 criteria, trained nurses, medical record analyst; (2) trained physicians	No§	Both	Yes	Before; after	HMPS like	No
Davis, 2002[42–44]	HMPS	Random	2-stage; (1) 18 criteria, trained nurses; (2) trained medical officers	No§	Both	Yes	Before; after	HMPS like	Yes

Continued

Table 2 Continued

First author, publication year	RRR method	Method of record selection	Method of review*	> 1 AE/ patient	Commission/ omission	Data on IRR outcome	Time frame of AE detection, in addition to index admission	AE definition†	Preventability reporting
Letaief, 2010[53]	HMPS	Random	2-stage; (1) 18 criteria, trained medical student; (2) physicians	No§	Both	Yes	Before; after	Wider than IHI	Yes
Requena, 2011[59]	HMPS	Random/total sample¶	2-stage; (1) 19 criteria, NS; (2) NS	Yes	NS	No	NS	Wider than IHI	Yes
Sommella, 2014[64]	HMPS	Random	2-stage; (1) 16 criteria, trained physicians; (2) trained physicians	No	NS	No	Before NS; after	HMPS like	No
Soop, 2009[65]	HMPS	Random	3-stage; (1) 18 criteria, trained nurses; (2) trained physicians; (3) member of the Scientific Council	No§	Both	Yes	Before; after	HMPS like	Yes
Wilson, 1995[12]	HMPS	Random	2-stage; (1) 18 criteria, trained nurses; (2) trained medical officers	No§	Both	Yes	Before; after	HMPS like	Yes
Woods, 2005[69 70 74]	HMPS	Random	2-stage; (1) 18 criteria, trained nurses; (2) trained physicians	No	Both	Yes	Before; after	HMPS like	Yes
Zegers, 2009[33 75]	HMPS	Random	3-stage; (1) 18 criteria, trained nurses; (2–3) trained physicians	Yes	Both	Yes	Before; after	HMPS like	Yes
Intensive care population, GTT/TT methodology**									
Agarwal, 2010[38]	TT	Random	2-stage; (1) 22 triggers, trained nurses, physicians; (2) trained physician, pharmacist	Yes	Both	No	Before; after	Wider than IHI	Yes
Barrionuevo, 2010[39]	TT	Total sample	2-stage; (1) 19 triggers, trained nurses; (2) trained physician	Yes	NS	No	NS	Wider than IHI	Yes
Hooper, 2014[48]	TT	Random	Unclear; 22 triggers, trained investigators	Yes	Commission only	Yes	Before; after	Wider than IHI	No
Jorro-Baron, 2021[49]	TT	Random	2-stage; (1) 37 triggers, trained PICU staff; (2) trained physicians	Yes	Both	No	NS	NS	Yes
Larsen, 2007[8]	TT	Every seventh admission	Unclear; 46 triggers, nurses, physicians	Yes	NS	Yes	NS	NS	Yes
Maziero, 2020[57]	TT	NS	2-stage; (1) NS, reviewer; (2) physician	Yes	NS	No	NS	NS	No
Sharek, 2006[62]	TT	Random	2-stage; (1) 17 triggers, trained nurses; (2) trained physician	Yes	Both	No	NS	Wider than IHI	Yes
Ventura, 2012[71]	TT	Total sample	Unclear; 14 triggers, researcher	Yes	NS	No	NS	Wider than IHI	Yes
Verlaat, 2018[72]	TT	Random	2-stage; (1) 19 triggers, trained physician; 2) trained physician	Yes	NS	Yes	Before; after	HMPS like	Yes

Continued

Table 2 Continued

First author, publication year	RRR method	Method of record selection	Method of review*	> 1 AE/ patient	Commission/ omission	Data on IRR outcome	Time frame of AE detection, in addition to index admission	AE definition†	Preventability reporting
Vermeulen, 2014[73]	TT	Random	2-stage; (1) 23 triggers, trained medical student; (2) medical director	Yes	Both	No	Before NS; after	Wider than IHI	No

IHI does not require any additional monitoring, treatment or hospitalisation.
*Number of review stage(s); number of triggers/screening criteria, record review training, type of reviewer(s)/review stage.
†HMPS like requires temporary or permanent disability, death or prolonged hospitalisation; IHI like requires additional monitoring, treatment or hospitalisation, or that results in death; Wider than IHI does not require any additional monitoring, treatment, or hospitalization.
‡Outcome for the total cohort, including both general and intensive care population.
§Additional data from authors.
¶This study merged paediatric data from three studies which used different sampling techniques.
**No study using the HMPS methodology included mainly intensive care patients.
AE, adverse event; GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study methodology; IHI, Institute for Healthcare Improvement; IRR, interrater reliability; NS, not specified; RRR, retrospective record review; TT, Trigger Tool.

Online supplemental material 1, figures S2–S5 present forest plots for the secondary outcomes. In 24 samples (GTT/TT), AEs per 100 admissions (general care, range 6.8–93.8; intensive care, 30.2–325.0) were supplied or could be calculated. The pooled estimate for the general care population was 24.8 AEs per 100 admissions (95% PI 4.2–145.2) and 103.6 AEs per 100 admissions (95% PI 5.3–699.7) for intensive care (table 4; online supplemental material 1, figure S2). An overview of the pooled estimates and related measures for the primary and secondary outcomes is shown in table 4.

In 22 samples (GTT/TT), AEs per 1000 patient days varied between 15.5 and 390.8 for general care and 22.6 and 599.1 for intensive care. The pooled estimates for AEs per 1000 patient days were 48.3 (95% PI 5.9–393.1) and 126.2 (95% PI 6.4–2495.1) for general care and intensive care, respectively. Half of the studies for intensive care had over 100 AEs per 1000 patient days (range 195.7–599.1) (table 4; online supplemental material, figure S3).

Of the 16 samples that reported preventability, the pooled percentage of preventable AEs for GTT/TT (general and intensive care populations) was 58.6% (95% PI 7.4%–96.2%) and 67.4% (95% PI 4.5%–98.9%). The corresponding for the HMPS was 53.2% (95% PI 10.4%–91.8%) (table 4; online supplemental material, figure S4). The pooled percentage of admissions with preventable AEs (12 samples) was for the GTT/TT (general and intensive care) 7.3% (95% PI 0.0%–100.0%) and 25.0% (95% PI 2.5%–81.3%) and for HMPS 2.3% (95% PI 0.0%–59.3%) (table 4; online supplemental material, figure S5).

Quality assessment and sensitivity analysis

Several methodological concerns were identified during the quality assessment process.

Concerning overall assessments, risk of bias was assessed as *high* in 85% compared with 44% in the 9 GTT/TT and 13 HMPS studies for the general population and 100% for the intensive care population (n=10, GTT/TT). When compared with GTT/TT studies, HMPS studies more frequently had both a *low* risk of bias with *low* applicability concerns at the domain level (online supplemental material 1, table S7, figures S6-S7).

The stratified analysis exploring heterogeneity was based on the quality assessment and percentage of admissions with ≥1 AE as the outcome. Lower AE outcomes were detected where the risk of bias was rated as *high* or *unclear* in the domain 'record review process' than in those with a *low* risk of bias for general care (GTT/TT) (online supplemental material 1, figure S8). For the HMPS methodology, variation is driven by the unclear category, which hampers interpretation (online supplemental material 1, figure S9). For the intensive care population, studies with *high* risk of bias detected lower levels of AEs in the domain 'patient

Table 3 Per cent of total number of adverse events (AEs) per type and study

Study reference number(s) → types of AEs	General care population									Intensive care population							
	GTT/TT methodology							HMPS methodology		GTT/TT methodology							
	40	41	45*	54–56	58	66	13 67	59	69 70 74	38	39	48	57	62	71	72	73
Nosocomial infection†	9.2	59.6	42.8	–	–	18.6	6.8	13.8	–	6.6	19.2	9.2	40.0	27.8	13.5	20.0	7.3
Pulmonary‡	3.6	1.9	–	–	4.7	12.8	16.8	–	–	19.2	13.7	36.7	16.7	11.0	10.5	17.8	23.2
Skin, tissue or blood vessel harm§	8.1	1.9	–	–	32.6	–	23.8	–	–	8.0	–	–	33.3	15.9	6.0	4.4	12.3
Medication related¶	2.5	3.8	48.6	12.5	2.3	–	–	37.9	19.1	–	–	1.0	3.3	–	–	–	–
Medical technical product (eg, catheter or tube)**	–	–	–	–	4.7	14.5	–	–	–	15.1	30.8	5.1	3.3	1.3	6.0	4.4	20.7
Gastrointestinal††	4.5	–	–	–	37.2	11.8	8.3	–	–	5.6	–	5.1	–	5.2	1.0	11.1	4.6
Neurological‡‡	1.0	–	–	–	–	5.3	3.3	–	–	4.7	10.3	2.0	–	16.0	31.1	13.3	5.1
Renal, endocrine, fluid and electrolytes§§	5.6	–	–	0.8	2.3	6.3	5.9	–	–	6.8	–	14.3	–	5.4	21.1	4.4	17.4
Surgical¶¶	0.9	9.6	–	30.5	–	8.7	7.5	–	16.3	1.7	–	7.1	–	0.2	–	–	1.9
Cardiovascular***	–	–	–	–	–	5.4	4.6	–	–	14.4	1.4	8.2	–	8.7	4.1	4.4	1.2
Haematological†††	3.4	1.9	–	–	4.7	5.1	3.8	–	–	3.9	12.3	2.0	3.3	1.1	0.5	6.7	1.2
Pain‡‡‡	–	–	–	–	11.6	–	5.8	–	–	9.8	–	–	–	–	–	–	1.9
Deterioration in vital signs§§§	7.5	1.9	–	–	–	–	–	–	–	–	–	2.0	–	–	–	11.1	–
Other¶¶¶	53.7	19.2	–	56.3	–	11.8	13.7	48.2	65.5	4.2	12.3	7.1	–	7.6	6.2	2.2	3.1

Examples of AEs within each type.
–Type not reported by the respective study.
*The only two AE types supplied in the publication.
†Central line-associated bloodstream infection, nosocomial infection, pneumonia, sepsis, urinary tract infection, wound infection.
‡Chylothorax, endotracheal tube malposition requiring repositioning, pneumothorax, postextubation stridor, reintubation, respiratory depression/compromise, unplanned extubation.
§Burn, blister, catheter infiltration, extravasation injury, nasal septum injury, pressure sore, skin breakdown, skin lesion, skin necrosis, skin problems.
¶Medication error, abrupt medication stop, drug level out of range, hyaluronic acid adverse reaction, medication related.
**Bladder catheter obstruction, feeding tube complication, intravenous catheter complication, tube complication (foley, chest drain or nasogastric tube).
††Abdominal compartment syndrome, antiemetic given, constipation, delay in diagnosis of gastric perforation, emesis/vomiting, necrosis of digits.
‡‡Abnormal cranial imaging, agitation/delirium, central nervous system bleed, delirium/agitation, neurological complication, oversedation, seizure, stroke, withdrawal symptoms.
§§Abnormal electrolyte levels, acute renal failure, blood glucose disorders, dehydration, fluid overload, hyperglycaemia, hypoglycaemia, renal dysfunction, urinary retention.
¶¶Removal/injury or repair of organ, return to surgery, surgical complication.
***Abnormal heart rate or blood pressure, arrhythmia, cardiac depression/compromise, cardiac rhythm derangements (eg, bradycardia, tachycardia, other arrhythmias), hypotension.
†††Anaemia unspecified, bleeding from feeding tube, blood transfusions, deep vein thrombosis, emboli, haemorrhage/haematoma, postoperative bleeding, thrombocytopenia.
‡‡‡Pain, postoperative pain, uncontrolled pain.
§§§Cardiac arrest/respiratory arrest, cardiac or pulmonary arrest, or rapid response team activation, resuscitation, vital sign changes.
¶¶¶Allergic reaction/hypersensitivity reaction, birth-related, blood sample redraws, care-related, complication of procedure or treatment, death, diagnostic error, fracture, other, readmission.
GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; TT, Trigger Tool.

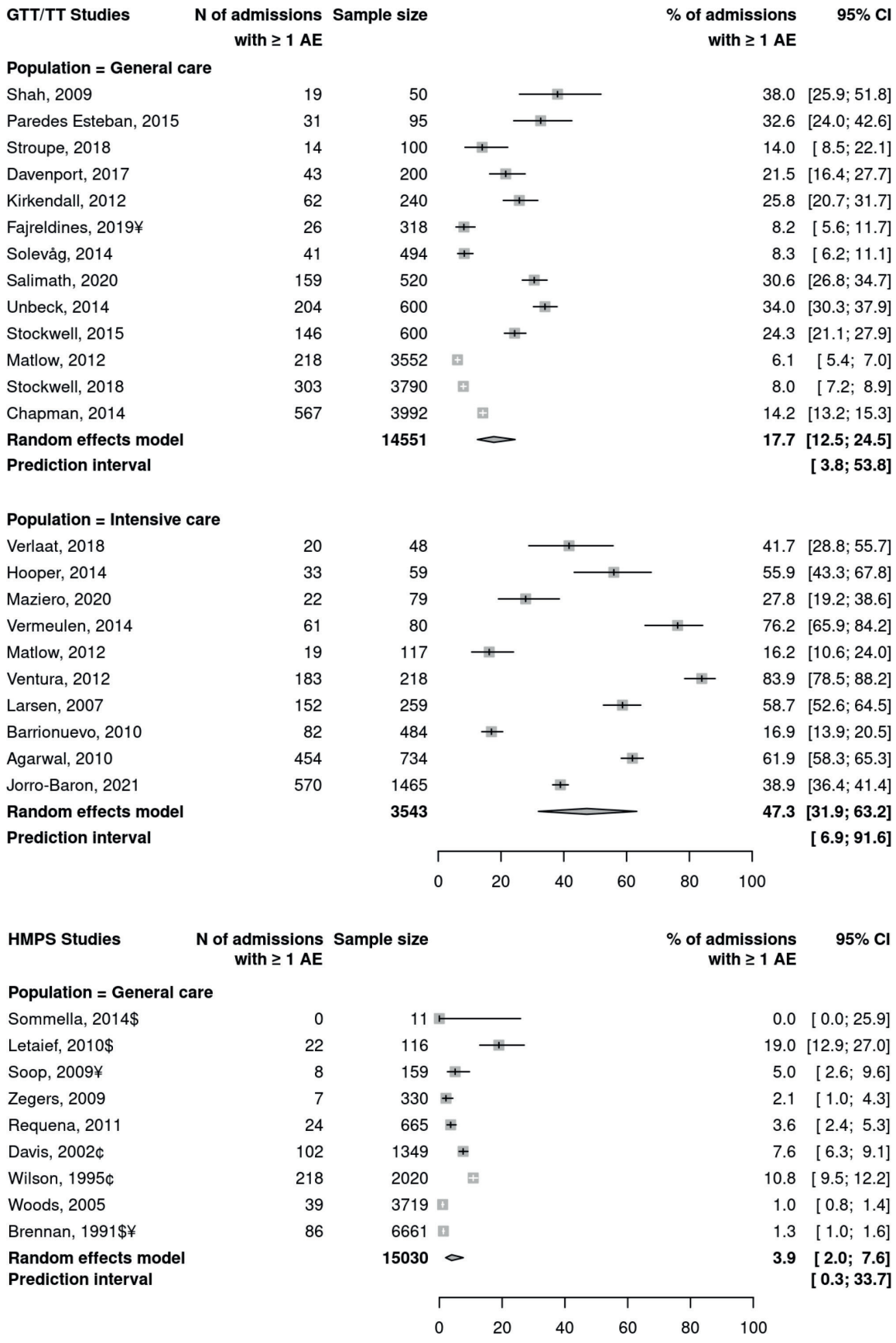


Figure 1 Forest plot of percentage of admissions with ≥1 adverse event (AE) for general care and intensive care populations and methodology, ordered by sample size. $ Sum of subgroups. ¥ Calculation of number of admissions with AEs. ¢ Scored 2–6 on the causation scale compared with 4–6 for other studies using this scale to determine whether an AE was caused by healthcare management rather than the patient's disease. GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; TT, Trigger Tool.

Table 4 Pooled estimates from meta-analyses

	General care population		General care population		Intensive care population		
	HMPS methodology		GTT/TT methodology				
	N of samples	Pooled estimates (95% PI)	N of samples	Pooled estimates (95% PI)	N of samples	Pooled estimates (95% PI)	P value between populations (GTT/TT)
Primary outcome							
Percentage of admissions with ≥1 AEs	9	3.9 (0.3–33.7)	13	17.7 (3.8–53.8)	10	47.3 (6.9–91.6)	0.0003
Secondary outcomes							
AEs per 100 admissions	–	–	13	24.8 (4.2–145.2)	11	103.6 (15.3–699.7)	< 0.0001
AEs per 1000 patient days	–	–	12	48.3 (5.9–393.1)	10	126.2 (6.4–2495.1)	0.0418
Percentage of preventable AEs	5	53.2 (10.4–91.8)	4	58.6 (7.4–96.2)	7	67.4 (4.5–98.9)	0.5355
Percentage of admissions with preventable AEs	4	2.3 (0.0–59.3)	3	7.3 (0.0–100.0)	5	25.0 (2.5–81.3)	0.0467

AEs, adverse events; GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; PI, prediction interval; TT, Trigger Tool.

selection' than those rated as *low* risk of bias (online supplemental material 1, figure S10). In all three strata, *high* risk of bias for the domain 'outcomes' was typically associated with higher AE rates compared with *low* risk of bias. Nevertheless, the limited sample size does not provide enough evidence to draw any solid conclusions.

DISCUSSION

We conducted a systematic review and a meta-analysis, consisting of 32 studies with 44 publications examining the incidence and characteristics of AEs detected using three commonly used record review methods (GTT, TT and HMPS). Nosocomial infections were common in both populations and most of the AEs were less severe. There was substantial between-study heterogeneity and overall high risk of bias in most studies. The PIs for the primary outcome for GTT/TT studies were 3.8%–53.8% and 6.9%–91.6% (general care and intensive care populations) and 0.3%–33.7% for the HMPS studies (general care). The PIs for the percentage of preventable AEs for GTT/TT studies were 7.4%–96.2% and 4.5%–98.9% (general care and intensive care) and the equivalent for HMPS studies was 10.4%–91.8%.

Incidence and characteristics of adverse events in paediatric inpatient care

Our review confirms substantial heterogeneity between general care and intensive care studies, as well as between methodologies (GTT/TT and HMPS). However, the results also display a high level of heterogeneity within populations. The degree of heterogeneity is in accordance with previously published systematic reviews.[22 23 25 76] The majority of studies were judged to be at high risk of bias, which also lowers the trust we place in the summary estimates. Therefore, caution is needed when drawing conclusions from the pooled data of the combined studies. We urge the reader to focus on the given PIs when interpreting the pooled data.

Berchialla *et al*[25] focused solely on paediatric inpatient AEs in their systematic review and reported a pooled incidence of AEs at 2.0%. This is lower than the pooled incidence of admissions with ≥1 AE using the GTT/TT methodology, shown in the present review: 17.7% for the general care population and 47.3% for the intensive care population. However, it is in line with the 3.9% for studies conducted using HMPS methodology. Their inclusion of studies using only the HMPS's AE definition may partly explain the difference, as the threshold for inclusion of an AE is higher due to the requirement of temporary or permanent disability, death or increased length stay. This could have led to minor, but perhaps commonly occurring AEs, being excluded with the risk of underestimation of AEs as a consequence.

Sauro *et al*[23] included, in addition to record review, other data collection methods, and found a pooled estimate of 1.4 AEs per 100 paediatric admissions and up to 11.9 AEs per 100 admissions in adult care. This is also considerably lower than our corresponding estimate for GTT/TT studies at 24.8 and 103.6 (general care and intensive care populations) AEs per 100 admissions. On the other hand, a newly published systematic review[29] including GTT/TT studies in general care reported a pooled incidence of 30.0 AEs per 100 adult admissions which is higher compared with our findings in the general care population.

Half of the studies assessed and reported preventability, both for general and intensive care, and around 60% of AEs were identified as preventable. However, as discussed by Hibbert *et al*,[6] the assessment of preventability is a subjective judgement and comparison between studies is to be done with caution. This, therefore, is a methodological limitation. Panagioti *et al*[22] included both adult and paediatric populations and showed an overall pooled prevalence of 6% for preventable AEs. This is in line with our preventability estimates in GTT/TT studies of 7.3% for admissions in general care but is higher compared with the 2.3% found in the HMPS studies. However, we found a pooled estimate of 25% for preventable AEs for intensive care compared with 18% in the study by Panagioti *et al*.[22]

A longitudinal retrospective record review study indicated an increased frequency of AEs over time, where one explanation was the increased number of patients with less complex conditions receiving day and outpatient care instead of inpatient care. This leads to an increased proportion of seriously ill patients in hospitals, and this may affect the AE rates for inpatient care.[77]

Important aspects of the variation in AE rates are the context and case mix of patients such as inclusion of units, medical specialities, hospital types, academic level of the hospital, patient age and comorbidity, and level of care. In both general care and intensive care populations, nosocomial infection was among the most common type of adverse event, also identified as one of the main causes of morbidity and mortality for paediatric inpatients.[78] Paediatric patients have many risk factors for infections related to, among other things, immunodeficiencies and poor skin barrier. Skin harm is a predisposing factor for nosocomial infections,[79] and was the overall third most common type of AEs in the current review. It is important to keep in mind the considerable variation regarding the taxonomy of reported types of AE used, which makes comparisons between studies difficult.

Study methodology

The use of record review methodology for specific populations seems to have increased over the last few decades. All studies conducted solely in the intensive care population were conducted after 2006 and a vast majority in the last 10 years.

We could not sufficiently explain the heterogeneity in the primary outcome using the quality of the studies. Insufficient reporting affected the risk of bias and applicability-related concerns negatively. The high risk of bias for the domain 'outcomes' was typically associated with a higher percentage of admissions with an AE. Sauro *et al*[23] reported, in accordance with our findings, a significantly higher pooled estimate of AEs for lower-quality studies. Furthermore, they showed, in consistency with the current study, that the presence of AEs at admission was unclear.

Many methodological limitations and reasons for the variations of AE outcomes in published studies have been suggested, for example, patient record documentation, the experience of the review team, quality assurance activities, inclusion criteria, AE definitions, choice of triggers and time frame for inclusion of AEs.[6] Apart from the researchers' adaptations, some variations may be explained by the different record review methods. Although, it would have been very interesting to analyse the variation based on the different methodological applications, it was outside of the scope of this review. In a recently published meta-analysis for adult inpatients, some of the variation could be explained by those methodological aspects (type of hospital included, age of sample included and experience of the review team).[29]

Another aspect is that variables that might affect the estimates of AE outcomes were not always clearly specified in the studies, for example, the time frames for AE inclusion. As a consequence, data extractors made interpretations based on triggers, for example, hospital readmission within 30 days. Another example is the inclusion of acts of commissions and/or omissions which was often not explicitly specified in the studies. GTT and TT studies following the Institute for Healthcare Improvement's manual exclude AEs related to acts of omission which could lead to an underestimation of AEs. Wilson *et al*[12] found in their study that acts of omission were nearly twice as common as acts of commission. Hibbert *et al*[6] suggest that several additional variables should be included when using GTT, for example, omissions, preventability and other characterisations, to get a better understanding of AEs. This suggestion is in accordance with the HMPS methodology, where AEs are categorised to a higher extent compared with GTT. To summarise, as many studies use minor adaptations of the record review process,[19 80] the reporting of AEs would benefit from a standardised guideline. This would decrease the methodological heterogeneity, thereby increasing replicability, interpretations and comparisons.

Clinical implications

Despite variations between inpatient care, AE outcomes and measurements, the high incidence of AEs and

percentage of preventable AEs indicate that there is more to be done regarding patient safety interventions. Zegers *et al*[81] made an umbrella review concerning evidence-based interventions to reduce inpatient AEs and they conclude a need for more high-quality studies to determine what interventions will have the most positive impact on patient safety. However, they state that there is evidence available for interventions to prevent infections, falls, delirium, adverse drug events, cardiopulmonary arrest and mortality. Furthermore, the measurement of AEs must be incorporated as part of the learning system within healthcare organisations and be connected to evidence-based interventions and evaluation of these as part of the continuous improvement work as measurement alone does not create safe care.[82]

Strengths and limitations

The adoption of a robust search strategy using several databases with no limitations in publication dates or language of publication lessens the likelihood that important studies were missed and may have changed the estimates in a significant way. However, the possibility of missing potentially relevant studies meeting the inclusion criteria is always present as we did not search for 'grey' literature. We did not use funnel plots to explore publication bias or other biases associated with small study size, as patterns of publication bias in the field of single-arm studies reporting proportions is not well understood and also because funnel plot analyses can lead to inaccurate conclusions.[83]A rigorous approach was adopted to the screening and data extraction process, as well as the assessment of bias and applicability. The large number of studies included further strengthens the study. We also contacted the authors for several of the studies where vital variables were missing. This led to fewer variables being categorised as not specified and therefore fewer studies were excluded from the meta-analyses.

One limitation is that the exclusion criteria disqualified studies with, for example, only automated AE detection, those including only outpatients or studies focusing on a specific diagnosis, treatment, or AE such as adverse drug events. This could have reduced the number of eligible studies and the final sample size as estimates could differ from estimates in a wider population. Concerning generalisability, most studies were conducted in Europe, as well as North and South America. Last but not the least, critically ill patients need complex care, which puts them at risk for AEs.[3] As previously stated, paediatric patients run a high risk for AEs during inpatient care, in general care, but specifically in intensive care.[7] Some of the heterogeneity within the general care population might be explained by the fact that several studies in the general care population also included intensive care patients to some extent. We choose to include a heterogeneous group of studies to provide estimates of paediatric inpatient AEs to represent the diversity of hospital settings, as well as to include the three most common record review methodologies.

For the reporting of the meta-analysis, we have taken the decision to not report on I^2 values. This measure can be used to compare statistical heterogeneity but not clinical heterogeneity.[84] Rücker *et al*[84] recommends using τ^2 to assess clinical heterogeneity. IntHout *et al*[34] go a step further and recommend presenting PIs instead, as it is presented on the same scale as the outcome measure in contrast to τ^2 or I^2. Therefore, we opted to provide PIs as measures of heterogeneity.

We acknowledge a deviation from the published study protocol, as we changed our primary outcome measure during the data-extraction phase, before conducting any statistical analyses. The percentage of admissions with ≥1 AE was chosen instead of AEs per 100 admissions, because this was the only measure with which we could directly compare the two methodological groups of GTT/TT and HMPS.

CONCLUSION

This review demonstrates a large between-study variation in estimates of the incidence of paediatric AEs. It also highlights the importance of a thorough understanding of the complex nature of AEs, and the sources of variation and of bias. The current lack of reporting standards in this field impedes comparison of study results. To advance the field of record review methodology, new reporting and risk of bias guidance tools are needed to enhance both comparability and overall quality of the studies and to maximise impact of study findings.

Author affiliations
[1]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden
[2]Division of Pediatrics, Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden
[3]Institute of Nursing Science, Department Public Health, Faculty of Medicine, University of Basel, Basel, Switzerland
[4]Department of Medical and Surgical Sciences SMECHIMAI, University of Modena and Reggio Emilia, Modena, Italy
[5]Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland
[6]School of Health and Welfare, Dalarna University, Falun, Sweden
[7]Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[8]Department of Mathematics and Computer Science, University of Basel, Basel, Switzerland
[9]Department of Clinical Sciences, Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden

Twitter Luisa C Eggenschwiler @l_eggens, Anne W S Rutjes @AnneRutjes and Michael Simon @msimoninfo

Acknowledgements The authors wish to thank the librarians Alena Lindfors, Ulrika Gabrielsson and Kristina Lönn at Dalarna University for support with the development and tests of search strategies as well as the performance of the database searches.The authors would also like to thank Paula Kelly-Pettersson for the editing of this manuscript. An acknowledgement also to the ones who provided supplemental data for the data analyses: Swati Agarwal, Ross Baker, Laura Barrionuevo, Troyen Brennan, Peter Davis, Sana El Mhamdi, María Eugenia Esandi, Virginia Flintoft, Robert Giddered,

Stephen Hancock, Facundo Jorro Barón, Alexander Kiss, Gitte Larsen, Roy Lay-Yee, Lucian Leape, Mondher Letaief, Anne Matlow, Geetanjali Shankarprasad Salimath, Stephan Schug, Paul Sharek, Anne Lee Solevåg, Michael Soop, Carin Verlaat and Jentien Vermeulen.

Contributors PD, LCE, AWSR, LB, SNM, MS, UF and MU contributed to conceptualisation of the project and to methodology. PD, LCE and MU carried out data curation. LCE performed formal analysis with support from AWSR and GM. PD, LCE, UF and MU wrote the original draft. All authors contributed to review and editing of the manuscript. PD and LCE contributed equally to this paper. Author responsible for the overall content is MU.

Funding This study was funded by grants from a regional agreement on clinical research (ALF) between Region Stockholm and Karolinska Institutet (2020-0443), Childhood Foundation of the Swedish Order of Freemasons (no award/grant number).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data and R code are available on 10.5281/zenodo.7335611.

ORCID iDs
Pernilla Dillner http://orcid.org/0000-0001-7350-7572
Luisa C Eggenschwiler http://orcid.org/0000-0002-1939-2345
Anne W S Rutjes http://orcid.org/0000-0001-9782-779X
Lena Berg http://orcid.org/0000-0003-1815-799X
Sarah N Musy http://orcid.org/0000-0003-0267-3375
Michael Simon http://orcid.org/0000-0003-2349-7219
Giusi Moffa http://orcid.org/0000-0002-2739-0454
Ulrika Förberg http://orcid.org/0000-0003-4291-425X
Maria Unbeck http://orcid.org/0000-0002-5090-0352

REFERENCES

1 Tessier L, Guilcher SJT, Bai YQ, *et al*. The impact of hospital harm on length of stay, costs of care and length of person-centred episodes of care: a retrospective cohort study. *CMAJ* 2019;191:E879–85.
2 Lanzillotti LdaS, De Seta MH, de Andrade CLT, *et al*. Adverse events and other incidents in neonatal intensive care units. *Cien Saude Colet* 2015;20:937–46.
3 Kohn LT, Corrigan JM, Donaldson MS. *To err is human: building a safer health system*. Washington (DC): Institute of Medicine Committee on Quality of Health Care in, America, 1999.
4 World Health Organisation. World Alliance for patient safety: forward programme 2005 France, 2004. Available: https://apps.who.int/iris/handle/10665/43072
5 Griffin F, Resar R. *IHI global trigger tool for measuring adverse events (second edition*. Cambridge, Massachusetts: Instutute for Healthcare Improvment, 2009.
6 Hibbert PD, Molloy CJ, Hooper TD, *et al*. The application of the global trigger tool: a systematic review. *Int J Qual Health Care* 2016;28:640–9.
7 Kalisch BJ, Landstrom G, Williams RA. Missed nursing care: errors of omission. *Nurs Outlook* 2009;57:3–9.
8 Larsen GY, Donaldson AE, Parker HB, *et al*. Preventable harm occurring to critically ill children. *Pediatr Crit Care Med* 2007;8:331–6.
9 Valentin A, Capuzzo M, Guidet B, *et al*. Patient safety in intensive care: results from the multinational sentinel events evaluation (see) study. *Intensive Care Med* 2006;32:1591–8.
10 Murff HJ, Patel VL, Hripcsak G, *et al*. Detecting adverse events for patient safety research: a review of current methodologies. *J Biomed Inform* 2003;36:131–43.
11 Brennan TA, Leape LL, Laird NM, *et al*. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard medical practice study I. *N Engl J Med* 1991;324:370–6.
12 Wilson RM, Runciman WB, Gibberd RW, *et al*. The quality in Australian health care study. *Med J Aust* 1995;163:458–71.
13 Landrigan CP, Stockwell D, Toomey SL, *et al*. Performance of the global assessment of pediatric patient safety (GAPPS) tool. *Pediatrics* 2016;137:06.
14 Unbeck M, Lindemalm S, Nydert P, *et al*. Validation of triggers and development of a pediatric trigger tool to identify adverse events. *BMC Health Serv Res* 2014;14:655.
15 Mattsson TO, Knudsen JL, Brixen K, *et al*. Does adding an appended oncology module to the global trigger tool increase its value? *Int J Qual Health Care* 2014;26:553–60.
16 Nilsson L, Borgstedt-Risberg M, Brunner C, *et al*. Adverse events in psychiatry: a national cohort study in Sweden with a unique psychiatric trigger tool. *BMC Psychiatry* 2020;20:44.
17 Lindblad M, Schildmeijer K, Nilsson L, *et al*. Development of a trigger tool to identify adverse events and no-harm incidents that affect patients admitted to home healthcare. *BMJ Qual Saf* 2018;27:502–11.
18 Classen DC, Resar R, Griffin F, *et al*. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff* 2011;30:581–9.
19 Klein DO, Rennenberg RJMW, Koopmans RP, *et al*. A systematic review of methods for medical record analysis to detect adverse events in hospitalized patients. *J Patient Saf* 2021;17:e1234–40.
20 Naessens JM, Campbell CR, Huddleston JM, *et al*. A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care* 2009;21:301–7.
21 de Vries EN, Ramrattan MA, Smorenburg SM, *et al*. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216–23.
22 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.

23 Sauro KM, Machan M, Whalen-Browne L, *et al*. Evolving factors in hospital safety: a systematic review and meta-analysis of hospital adverse events. *J Patient Saf* 2021;17:e1285–95.
24 Schwendimann R, Blatter C, Dhaini S, *et al*. The occurrence, types, consequences and preventability of in-hospital adverse events - a scoping review. *BMC Health Serv Res* 2018;18:521.
25 Berchialla P, Scaioli G, Passi S, *et al*. Adverse events in hospitalized paediatric patients: a systematic review and a meta-regression analysis. *J Eval Clin Pract* 2014;20:551–8.
26 Bramer WM, Rethlefsen ML, Kleijnen J, *et al*. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev* 2017;6:245.
27 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
28 Musy SN, Ausserhofer D, Schwendimann R, *et al*. Trigger Tool-Based automated adverse event detection in electronic health records: systematic review. *J Med Internet Res* 2018;20:e198.
29 Eggenschwiler LC, Rutjes AWS, Musy SN, *et al*. Variation in detected adverse events using trigger tools: a systematic review and meta-analysis. *PLoS One* 2022;17:e0273800.
30 R Core Team. R:A Language and Environment for Statistical Computing Vienna, Austria: R Foundation for Statistical Computing, 2022. Available: https://www.R-project.org/
31 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.
32 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *JSS* 2010;36:1–48.
33 Zegers M, de Bruijne 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.
34 IntHout J, Ioannidis JPA, Rovers MM, *et al*. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016;6:e010247.
35 Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:d549.
36 Harbord RM, Deeks JJ, Egger M, *et al*. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;8:239–51.
37 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
38 Agarwal S, Classen D, Larsen G, *et al*. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med* 2010;11:568–78.
39 Barrionuevo LS, Esandi ME. [Epidemiology of adverse events in the neonatal unit of a regional public hospital in Argentina]. *Arch Argent Pediatr* 2010;108:303–10.
40 Chapman SM, Fitzsimons J, Davey N, *et al*. Prevalence and severity of patient harm in a sample of UK-hospitalised children detected by the paediatric trigger tool. *BMJ Open* 2014;4:e005066.
41 Davenport MC, Domínguez PA, Ferreira JP, *et al*. Measuring adverse events in pediatric inpatients with the global trigger tool. *Arch Argent Pediatr* 2017;115:357–63.
42 Davis P, Lay-Yee R, Briant R, *et al*. Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J* 2002;115:U271.
43 Davis P, Lay-Yee R, Briant R, *et al*. Adverse events in New Zealand public hospitals II: preventability and clinical context. *N Z Med J* 2003;116:U624.
44 Davis P, Lay-Yee R, Briant R. *Adverse events on new Zealand public hospitals: principal findings from a national survey*. Occasional Paper No 3: Wellington: Ministry of Health, 2001.
45 Fajreldines A, Schnitzler E, Torres S, *et al*. Measurement of the incidence of care-associated adverse events at the Department of pediatrics of a teaching hospital. *Arch Argent Pediatr* 2019;117:e106–9.
46 Geetanjali S, Ganachari MS. Development of a pediatric focused trigger tool to assess the prevalence of adverse events at a hospital setting: a retrospective structured chart review. *IJPR* 2020;12:1387–98.
47 Hiatt HH, Barnes BA, Brennan TA, *et al*. A study of medical injury and medical malpractice. *N Engl J Med* 1989;321:480–4.
48 Hooper AJ, Tibballs J. Comparison of a trigger tool and voluntary reporting to identify adverse events in a paediatric intensive care unit. *Anaesth Intensive Care* 2014;42:199–206.
49 Jorro-Barón F, Suarez-Anzorena I, Burgos-Pratx R, *et al*. Handoff improvement and adverse event reduction programme implementation in paediatric intensive care units in Argentina: a stepped-wedge trial. *BMJ Qual Saf* 2021;30:782–91.
50 Kirkendall ES, Kloppenborg E, Papp J, *et al*. Measuring adverse events and levels of harm in pediatric inpatients with the global trigger tool. *Pediatrics* 2012;130:e1206–14.
51 Lander L, Roberson DW, Plummer KM, *et al*. A trigger tool fails to identify serious errors and adverse events in pediatric otolaryngology. *Otolaryngol Head Neck Surg* 2010;143:480–6.
52 Leape LL, Brennan TA, Laird N, *et al*. The nature of adverse events in hospitalized patients. Results of the Harvard medical practice study II. *N Engl J Med* 1991;324:377–84.
53 Letaief M, El Mhamdi S, El-Asady R, *et al*. Adverse events in a Tunisian Hospital: results of a retrospective cohort study. *Int J Qual Health Care* 2010;22:380–5.
54 Matlow A, Flintoft V, Orrbine E, *et al*. The development of the Canadian paediatric trigger tool for identifying potential adverse events. *Healthc Q* 2005;8 Spec No:90–3.
55 Matlow AG, Baker GR, Flintoft V, *et al*. Adverse events among children in Canadian hospitals: the Canadian paediatric adverse events study. *CMAJ* 2012;184:E709–18.
56 Matlow AG, Cronin CMG, Flintoft V, *et al*. Description of the development and validation of the Canadian paediatric trigger tool. *BMJ Qual Saf* 2011;20:416–23.
57 Maziero ECS, Cruz EDdeA, Alpendre FT, *et al*. Association between nursing work conditions and adverse events in neonatal and pediatric intensive care units. *Rev Esc Enferm USP* 2020;54:e03623.
58 Paredes Esteban RM, Garrido Pérez JI, Ruiz Palomino A, *et al*. [Implementation of a Plan of Patient Safety in Service of Pediatric Surgery. First results]. *Cir Pediatr* 2015;28:111–7.
59 Requena J, Miralles JJ, Mollar J. Clinical safety paediatric patients. *RCA* 2011;26:353–8.
60 Salimath GS, Ganachari MS, Gudhoor M. Paediatric focused triggering tool (PFTT) to assess the harm and its utilization to minimize the levels of harm among children at a tertiary care hospital. *IJPER* 2020;54:819–25.
61 Shah RK, Lander L, Forbes P, *et al*. Safety on an inpatient pediatric otolaryngology service: many small errors, few adverse events. *Laryngoscope* 2009;119:871–9.

62 Sharek PJ, Horbar JD, Mason W, *et al*. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics* 2006;118:1332–40.
63 Solevåg AL, Nakstad B. Utility of a paediatric trigger tool in a Norwegian department of paediatric and adolescent medicine. *BMJ Open* 2014;4:e005011.
64 Sommella L, de Waure C, Ferriero AM, *et al*. The incidence of adverse events in an Italian acute care Hospital: findings of a two-stage method in a retrospective cohort study. *BMC Health Serv Res* 2014;14:358.
65 Soop M, Fryksmark U, Köster M, *et al*. The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. *Int J Qual Health Care* 2009;21:285–91.
66 Stockwell DC, Bisarya H, Classen DC, *et al*. A trigger tool to detect harm in pediatric inpatient settings. *Pediatrics* 2015;135:1036–42.
67 Stockwell DC, Landrigan CP, Toomey SL, *et al*. Adverse events in hospitalized pediatric patients. *Pediatrics* 2018;142:08.
68 Stroupe LM, Patra KP, Dai Z, *et al*. Measuring harm in hospitalized children via a trigger tool. *J Pediatr Nurs* 2018;41:9–15.
69 Thomas EJ, Orav EJ, Brennan TA. Hospital ownership and preventable adverse events. *Int J Health Serv* 2000;30:745–61.
70 Thomas EJ, Studdert DM, Burstin HR, *et al*. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000;38:261–71.
71 Ventura CMU, Alves JGB, Meneses JdoA. [Adverse events in a Neonatal Intensive Care Unit]. *Rev Bras Enferm* 2012;65:49–55.
72 Verlaat CW, van der Starre C, Hazelzet JA, *et al*. The occurrence of adverse events in low-risk non-survivors in pediatric intensive care patients: an exploratory study. *Eur J Pediatr* 2018;177:1351–8.
73 Vermeulen JM, van Dijk M, van der Starre C, *et al*. Patient safety in South Africa: PICU adverse event registration*. *Pediatr Crit Care Med* 2014;15:464–70.
74 Woods D, Thomas E, Holl J, *et al*. Adverse events and preventable adverse events in children. *Pediatrics* 2005;115:155–60.
75 Zegers M, de Bruijne MC, Wagner C, *et al*. Design of a retrospective patient record study on the occurrence of adverse events among patients in Dutch hospitals. *BMC Health Serv Res* 2007;7:27.
76 Connolly W, Li B, Conroy R, *et al*. National and institutional trends in adverse events over time: a systematic review and meta-analysis of longitudinal retrospective patient record review studies. *J Patient Saf* 2021;17:141–8.
77 Baines RJ, Langelaan M, de Bruijne MC, *et al*. Changes in adverse event rates in hospitals over time: a longitudinal retrospective patient record review study. *BMJ Qual Saf* 2013;22:290–8.
78 Morillo-García Áurea, Aldana-Espinal JM, Olry de Labry-Lima A, *et al*. Hospital costs associated with nosocomial infections in a pediatric intensive care unit. *Gac Sanit* 2015;29:282–7.
79 Tang YH, Jeng MJ, Wang HH. Risk factors and predictive markers for early and late onset neonatal bacteremic sepsis in preterm and term infants. *J Chin Med Assoc* 2021.
80 Doupi P, Svaar H, Bjørn B, *et al*. Use of the global trigger tool in patient safety improvement efforts: Nordic experiences. *Cogn Technol Work* 2015;17:45–54.
81 Zegers M, Hesselink G, Geense W, *et al*. Evidence-Based interventions to reduce adverse events in hospitals: a systematic review of systematic reviews. *BMJ Open* 2016;6:e012555.
82 Sauro K, Ghali WA, Stelfox HT. Measuring safety of healthcare: an exercise in futility? *BMJ Qual Saf* 2020;29:341–4.
83 Hunter JP, Saratzis A, Sutton AJ, *et al*. In meta-analyses of proportion studies, funnel plots were found to be an inaccurate method of assessing publication bias. *J Clin Epidemiol* 2014;67:897–903.
84 Rücker G, Schwarzer G, Carpenter JR, *et al*. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008;8:79.

Supplementary Material 1

Incidence and characteristics of adverse events in paediatric inpatient care: a systematic review and meta-analyses

Contents

Databases, search strategies and outcomes

The search strategy was constructed, based on the search query with the eligibility criteria. The final search strategy was constructed and validated in MEDLINE. The last author compiled a test set of relevant publications, the final search strategy was constructed to identify at least 95% of these citations. The search strategy in MEDLINE was then translated for use in the remaining databases and was also validated against the test set. All steps were carried out by information specialists except for the EMBASE database search which was performed by AWSR. Duplicates were excluded in accordance with Bramer et al.[1] de-duplication method for EndNote. The search strategies and their outcomes for Medline, Embase, Web of Science and Google Scholar are presented in tables S1-S4.

Table S1 Ovid Medline database search, ALL 1946 to 2nd June 2021, search date 4th June 2021

Search terms		Items found
Concept		
1.	exp Iatrogenic Disease/	78 011
2.	exp Medical Errors/	116 321
3.	exp Patient Harm/	182
4.	adverse event*.tw.	175 944
5.	harm.tw.	51 945
6.	trigger*.tw.	304 989
7.	or/1-6	715 698
Population		
8.	exp Adolescent/	2 094 701
9.	exp Child/	1 972 385
10.	exp Infant/	1 170 419
11.	p?ediatric*.tw.	377 272
12.	neonat*.tw.	277 105
13.	child*.tw.	1 452 803
14.	newborn*.tw.	167 883
15.	infant*.tw.	420 148
16.	adolescen*.tw.	296 962
17.	premature*.tw.	135 272
18.	preschool.tw.	25 704
19.	teenager*.tw.	14 859
20.	or/8-19	4 421 231
Context		
21.	exp Hospitals/	285 846
22.	exp Inpatients/	23 916
23.	exp Hospitalization/	257 553
24.	exp Hospital Units/	118 753
25.	exp Hospital Departments/	190 887
26.	hospital*.tw.	1 374 930
27.	intensive care.tw.	156 013
28.	inpatient*.tw.	116 945
29.	or/21-28	1 815 773
Types of evidence source		
30.	(review* adj5 (record* or chart*)).tw.	124 371
31.	trigger tool.tw.	287
32.	Harvard Medical Practice*.tw.	26
33.	or/30-32	124 533
Combined sets / Limits: publication year, language		
34.	7 and 20 and 29 and 33	1389

Table S2 Ovid Embase database search, 1974 to 7th June 2021, search date 8th June 2021

Search Terms		Items found
Concept		
1.	exp iatrogenic disease/	888 950
2.	exp medical error/	147 543
3.	exp patient harm/	2122
4.	adverse event*.tw.	314 703
5.	harm.tw.	69 071
6.	trigger*.tw.	390 929
7.	or/1-6	1 663 470
Population		
8.	exp adolescent/	1 590 228
9.	exp infant/	1 024 663
10.	exp child/	2 741 828
11.	p?ediatric*.tw.	589 425
12.	neonat*.tw.	360 500
13.	child*.tw.	1 825 708
14.	newborn*.tw.	196 148
15.	infant*.tw.	480 946
16.	adolescen*.tw.	387 560
17.	premature*.tw.	177 034
18.	preschool.tw.	30 071
19.	teenager*.tw.	20 975
20.	or/8-19	4 416 514
Context		
21.	exp hospital/	1 202 608
22.	exp hospital patient/	195 246
23.	exp hospitalization/	411 147
24.	exp child hospitalization/	11 187
25.	exp "hospital subdivisions and components"/	586 726
26.	exp hospital department/	23 929
27.	hospital*.tw.	2 105 252
28.	intensive care.tw.	228 605
29.	inpatient*.tw.	197 436
30.	exp hospital infection/	48 667
31.	or/21-30	2 884 244
Types of evidene source		
32.	(review* adj5 (record* or chart*)).tw.	228 720
33.	trigger tool.tw.	449
34.	Harvard Medical Practice*.tw.	35
35.	or/32-34	228 986
Publication type		
36.	Conference Abstract.pt.	4 102 339
37.	Editorial.pt.	692 899
38.	or/37-38	4 795 238
Combined sets		
39.	and/7,20,31,35	3295
40.	39 not 38	1658

Table S3 Web of Science database search, indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years, search date 4th June 2021

Search Terms		Items found
Concept		
1.	TS=(iatrogenic)	26 603
2.	TS=("adverse event*")	175 510
3.	TS=(harm)	100 701
4.	TS=(trigger*)	450 002
5.	#4 OR #3 OR #2 OR #1	745 154
Population		
6.	TS=(p$ediatric*)	427 520
7.	TS=(neonat*)	283 334
8.	TS=(child*)	1 903 734
9.	TS=(newborn*)	152 320
10.	TS=(infant*)	449 274
11.	TS=(adolescen*)	484 521
12.	TS=(premature*)	151 696
13.	TS=(preschool)	52 943
14.	TS=(teenager*)	20 423
15.	#14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6	2 894 445
Context		
16.	TS=(hospital*)	1 179 674
17.	TS=("intensive care")	170 727
18.	TS=(inpatient*)	122 042
19.	#18 OR #17 OR #16	1 339 227
Types of evidene source		
20.	TS=(review* NEAR/4 (record* OR chart*))	111 024
21.	TS=("trigger tool")	406
22.	TS=("Harvard Medical Practice*")	83
23.	#22 OR #21 OR #20	111 349
Combined sets / Limits: publication year, language		
24.	#23 AND #19 AND #15 AND #5	643

Table S4 Google Scholar database search, (include citations), the first 100 publications were included in the screening process, search date 4th June 2021

Search string	Items found
((iatrogenic OR "adverse event*" OR harm OR trigger*) AND (pediatric* OR paediatric* OR neonat* OR child* OR newborn* OR infant* OR adolescen* OR premature* OR preschool OR teenager*) AND (hospital* OR "intensive care" OR inpatient*) AND ("trigger tool" OR "Harvard Medical Practice*"))	2290

Eligibility criteria

The inclusion and exclusion criteria for publications are shown in table S5.

Table S5 Eligibility criteria in hierarchical order

	Inclusion	Exclusion
Population	- Children, all age groups, if they have been cared for at a paediatric inpatient unit - Studies addressing both adults and children if data provided for children are reported separately	- Studies reporting adverse events (harm) for children with a specific disease, diagnosis, for example, rheumatoid arthritis or children undergoing specific treatments or procedures as intubation, x-ray as well as only deceased patients
Context	- Hospitalized patients, acute care settings, both acute and elective admissions - All levels of inpatient care - All types of specialties	- Primary care, psychiatric care, day care/ ambulatory care - Emergency departments or other outpatient units at the hospital
Types of evidence source	- Peer reviewed full text primary studies, reporting relevant quantitative outcome data - Applied manual retrospective medical record review using Global Trigger Tool, Trigger Tool or Harvard Medical Practice Study methodologies as data collection methods - No restriction in language - No restriction in publication years	- Study protocols with no AE outcome published - Conference abstracts and editorials - Systematic reviews - Studies using, for example, clinical incident reporting systems as the primary data source and later these incident reports are analysed using record review
Concept	- All studies irrespective of which adverse event definition was used	- Studies reporting only specific adverse events, for example, adverse drug event

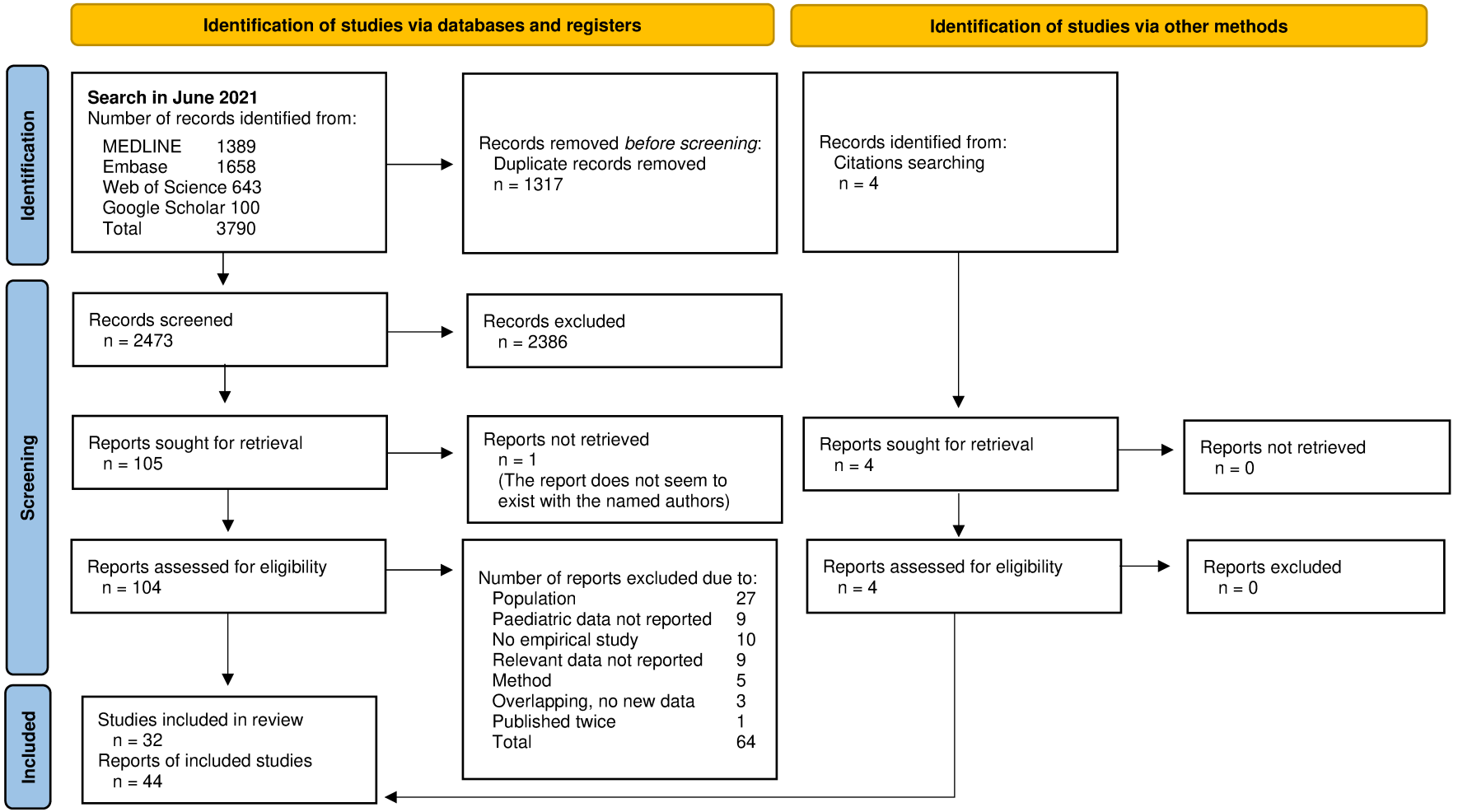


Figure S1 Flow diagram of literature search and included studies.[2]

Table S6 Adverse event (AE) outcomes and characteristics

First author, publication year	Admissions with AE n (%)	Admissions with preventable AE, n (%)	AEs/ 100 admissions	AEs/ 1000 patient days	Number of AEs range	Number of preventable AEs (%) range	Reporting type of AEs	Severity outcome of AEs
General care population, GTT/TT methodology								
Chapman, 2014 [3]	567 (14.2)	NA	25.1†	NS	1001, 0–10	NA	Yes	Minor, 60.4%; major, 31.8%; permanent, 1.8%; life threatening, 4.3%; death, 1.7%¶
Davenport, 2017 [4]	43 (21.5†)	NA	26.0	30.8	52, 0–8	NA	Yes	Minor, 40.0%; major, 46.0%; permanent, 1.9%; life threatening, 9,6%; death, 1,9%¶
Fajreldines, 2019 [5]	26 (8.2)	NA	11.0	15.5	35, NS	NA	Yes	Minor, 54.3%; major, 40.0%; permanent, 0%; life threatening, 5.7%; death, 0%¶
Kirkendall, 2012 [6]	62 (25.8)	NA	36.7	73.0†	88, NS	NA	No	Minor,76%; moderate, 22%; permanent, 0%; life threatening, 2%; death, 0%¶
Matlow, 2012 [7-9] (general care)	218 (6.1†)	99 (2.8†)	6.8†	16.4†	242, NS	NS	Yes	Minor, 16.5%; major, 62.7%; permanent, 6.1%; life threatening, 9.0%; death, 2.5%¶ Φ
Paredes Esteban, 2015 [10]	31 (32.6)	NS	45.3†	105.9	43, 0-5	NA	Yes	Minor, 88.4%; major, 11.6%; permanent, 0%; life threatening, 0%; death, 0%¶
Salimath, 2020 [11 12 19]	159§ (30.6†)	NS	35.0	66.4	182, NS	NA	No	Minor, 58.8%; major, 23.6%; permanent, 12.1%; life threatening, 5%; death, 0,6%¶
Shah, 2009 [13 14]	19 (38.0)	NS	68†	390.8 †	34, NS	NA	Yes	Minor, 76.5%; major, 23.5%; permanent, 0.0%; life threatening, 0.0%; death, 0.0%¶
Solevåg, 2014 [15]	41 (8.3†)	NS	8.3†	20.5†	41, 0–1	NA	No	NS separately for inpatients
Stockwell, 2015 [16]	146 (24.3)	NS	40.0	54.9	240 (NS)	108 (45.0), NS	Yes	Minor, 68.5%; major, 20.4%; permanent, 0.4%; life threatening, 10%; death, 0,4%¶
Stockwell, 2018 [17 18]	303† (8.0)	170† (4.5)	10.9	19.0	414, NS	210 (50.7), NS	Yes	Minor, 52.7%; major, 35.3%; permanent, 1.2%; life threatening, 10.1%; death, 0.7%¶
Stroupe, 2018 [19]	14 (14.0†)	NS	20.0	48.7	20, 0–3	11 (55.0), NS	No	Minor,60%; major, 35%; permanent, 0%; life threatening, 0%; death, 5%†¶
Unbeck, 2014 [20]	204 (34.0)	161 (26.8)	93.8†	101.3†	563, 0–27	442 (78.5), 0–22	No	NS
General care population, HMPS methodology								
Brennan, 1991 [21-23]	86 (1.3†)	NA	1.3†	NS	86, 0–1	NA	No	NS for paediatric population
Davis, 2002 [24-26]	102* (7.6†)	29*§ (2.1†)	7.6*	24.7†*§	102, 0–1	29 (28.4), 0–1	No	Permanent disability€ or death for 0-14 years, 6.9%; permanent disability€ or death for new-born/neonates, 2.9%
Letaief, 2010 [27]	22 (19.0†)	NS	19†	NS	22, 0–1	NS	No	NS for paediatric population
Requena, 2011 [28]	24 (3.6)	NS	4.4†	8.7	29, NS	19 (65.5), NS	Yes	Minor, 31.0%; moderate, 55.2%; severe, 0.0%

Sommella, 2014 [29]	0 (0.0†)	NS	0.0†	NS	0, 0–0	NA	No	NA
Soop, 2009 [30]	8 (5.0†)	4.4	5.0	NS	8, 0–1	7 (87.5), 0–1	No	NS for paediatric population
Wilson, 1995 [31]	218*§ (10.8†)	106*§ (5.2†)	10.8*	25.1*	218, 0–1	106 (48.6), 0–1	No	Minimal impairment×, 7.1%; moderate impairment‡, NS; permanent disability€, NS; death, 0%; unclear, NS
Woods, 2005 [32-34]	39 (1.0)	22† (0.6)	1.0†	NS	39, 0–1	22 (56.4), NS	Yes	NA
Zegers, 2009 [35 36]	7 (2.1†)	NS	NS	NS	NS	NS	No	NS
Intensive care population#								
Agarwal, 2010 [37]	454 (61.9)	278§ (37.9†)	202.7†	286.1†	1488, 0–34	683 (45.9), 0-20	Yes	Major + severe + death, 27%; severe + death, 10.0%
Barrionuevo, 2010 [38]	82 (16,9)	NS	30.2†	22.6†	146, 0–8	142 (97.3), NS	Yes	Without risk of patient's death, 65%; death, 5.6%†
Hooper, 2014 [39]	33 (55.9)	NA	166.1†	599.1†	98, NS	NA	Yes	Insignificant, 21% and 15%; minor, 32% and 36%; moderate, 35% and 37%; major, 7% and 9%; catastrophic, 4% and 3%
Jorro-Baron, 2021 [40]	570§ (38.9†)	444§ (30.3†)	93.8†	86.7†	1374, NS	992 (72.2), NS	No	NS
Larsen, 2007 [41]	152 (58.7)	88 (34.0)	195.8†	527†	507, NS	183 (36,1), NS	No	Minor patient harm, 80%; moderate patient harm, 17%; serious harm, 3%
Matlow, 2012 [7-9] (ICU cohort)	19 (16.2†)	7 (6,0†)	31.6†	23.5†	37, NS	NS	No	NS
Maziero, 2020 [42]	22 (27.8†)	NA	38.0†	NS	30, NS	NA	Yes	NA
Sharek, 2006 [43]	NS	NS	74.0	32.4	554, 0–11	312 (56.3), NS	Yes	Minor,60%; major, 17.3%; permanent, 6.5%; life threatening, 6.5%; death, 9.7%¶
Ventura, 2012 [44]	183 (83.9†)	NS	265.6†	195.7	579, NS	504 (87.0), NS	Yes	Minor, 40.0%; major, 36.0%; permanent, 14.8%; life threatening, 9.2%; death, 0.0%¶
Verlaat, 2018 [45]	20 (41.7†)	13 (27.1†)	93.7†	74†	45, 0–7	15 (33.3), NS	Yes	Minor, 44.4%; major, 0.0%; permanent, 11.1%; life threatening, 28.9%; death 15.6%†¶
Vermeulen, 2014 [46]	61 (76.2†)	NA	325.0†	507.8	260, NS	NA	Yes	NA

† Calculations are made.
* Scored 2–6 on the causation scale compared to 4–6 for other studies using this scale to determine whether an AE was caused by healthcare management rather than the patient´s disease.
§ Additional data from authors.
× Minimal impairment, temporary disability < 1 month.
‡ Moderate impairment, temporary disability 1-12 months.
€ Permanent disability: Davis et al. included <50% of function and > 50% of function; Wilson et al. included AEs which caused permanent impairment, or which resulted in permanent institutional or nursing care or death.
Φ Outcome for total cohort and the NCC MERP scale as two severity scales were used.
Studies using the HMPS methodology did not predominantly include intensive care patients.
¶ Modified version of the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) scale according to the Institute for Healthcare Improvement's manual for the Global Trigger Tool, that classifies level of harm from category E to I where E represents "temporary harm to the patient and required intervention"; F "temporary harm to the patient and required initial or prolonged hospitalization"; G "permanent patient harm", H "intervention required to sustain life"; I "patient death". NCC MERP E-I is categorised, due to the studies´ individual interpretation of the scale as: E= minor, F= major, G= permanent, H= life threatening, I= contributed to or resulted in patients' death.
AE, adverse event; GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; ICU, intensive care unit; NA, not applicable, did not include preventability and/or severity, NS, not specified; TT, Trigger Tool.

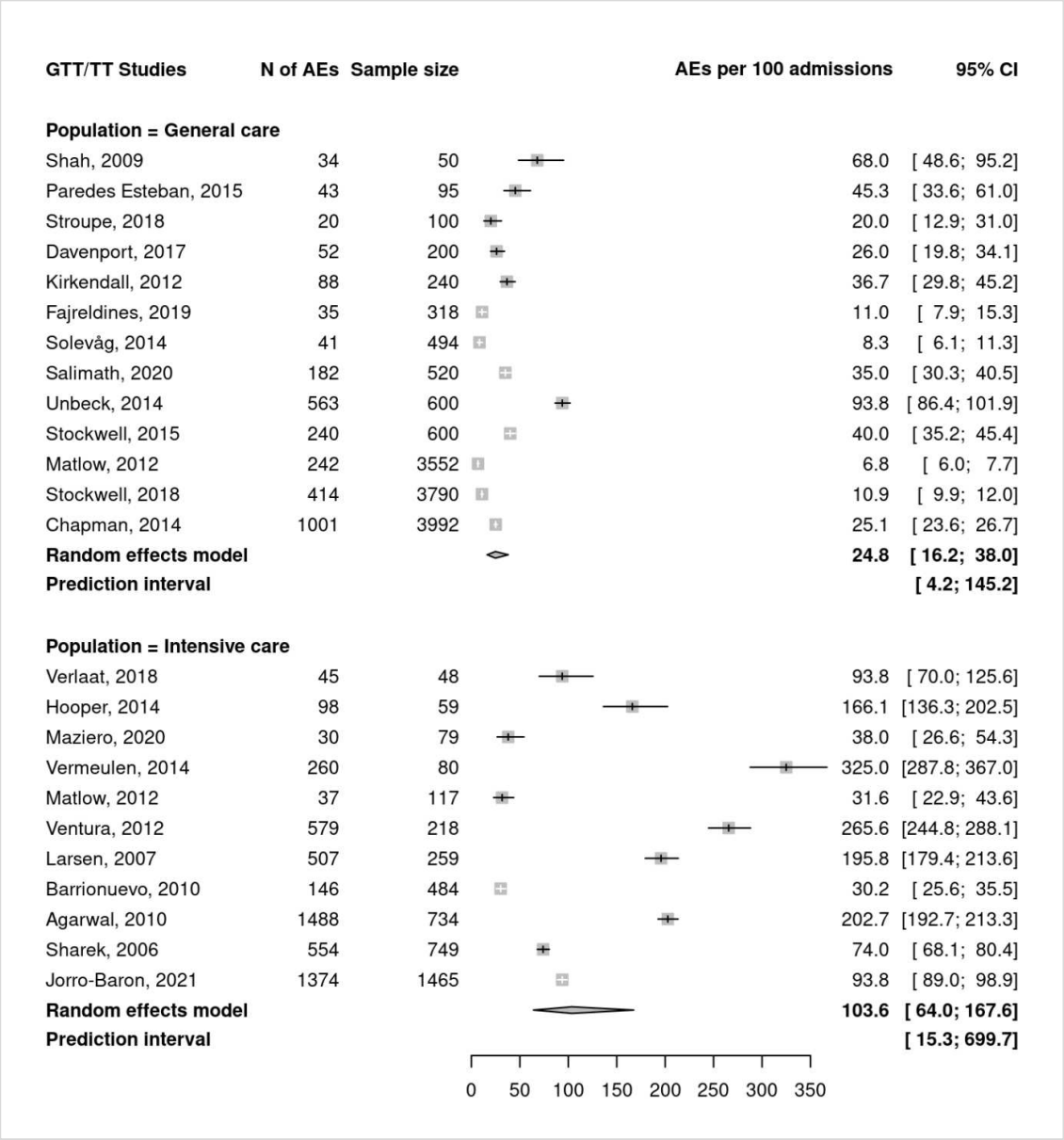


Figure S2 Forest plot of adverse events (AEs) per 100 admissions by general care respective intensive care population, ordered by sample size (GTT/TT). [3-20 28 37-46]
GTT, Global Trigger Tool; TT, Trigger Tool.

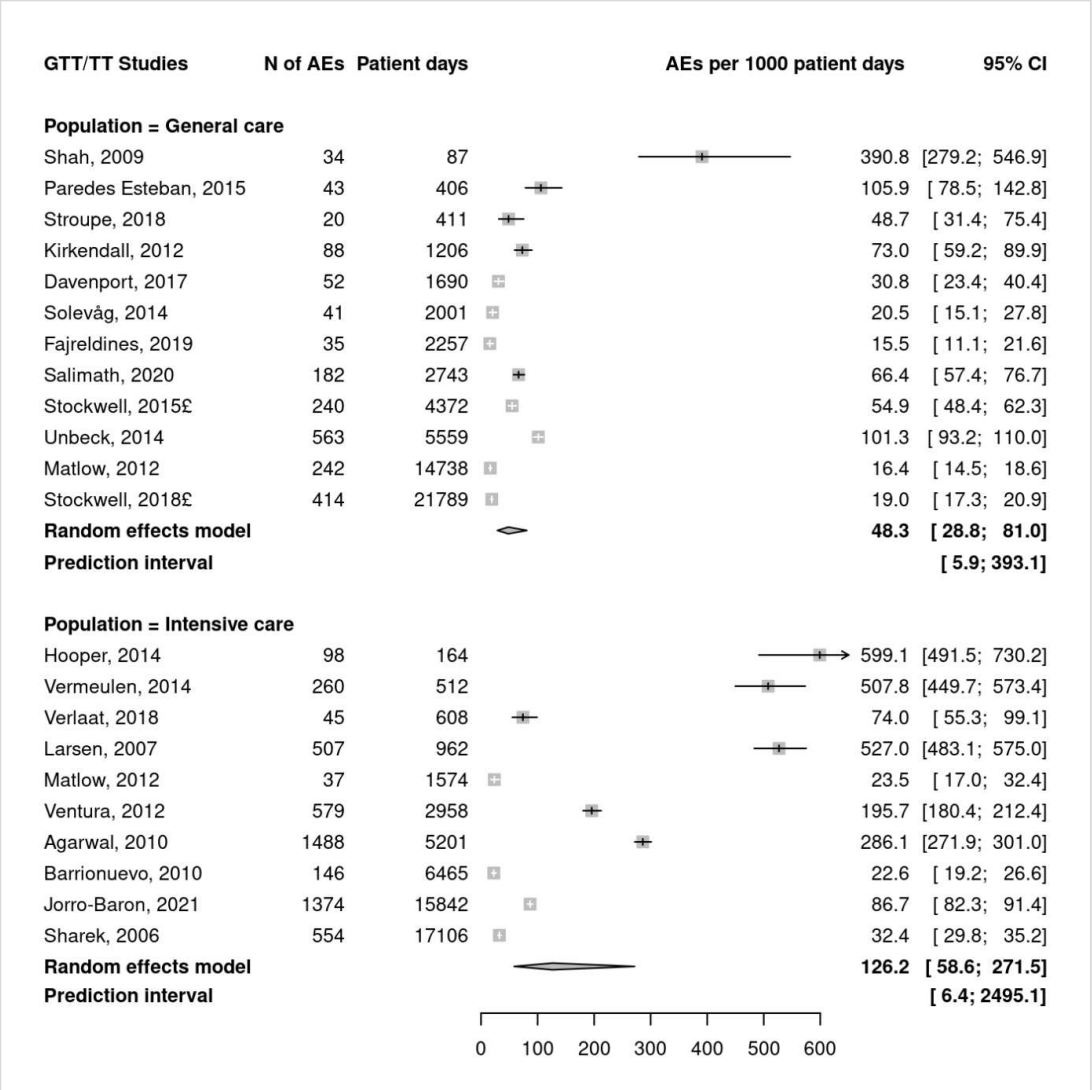


Figure S3 Forest plot of adverse events (AEs) per 1000 patient days by general care respective intensive care population, ordered by number of patient days (GTT/TT). [4-20 27 28 37-41 43-46]
£ Calculation of total number of hospital days.
GTT, Global Trigger Tool; TT, Trigger Tool.

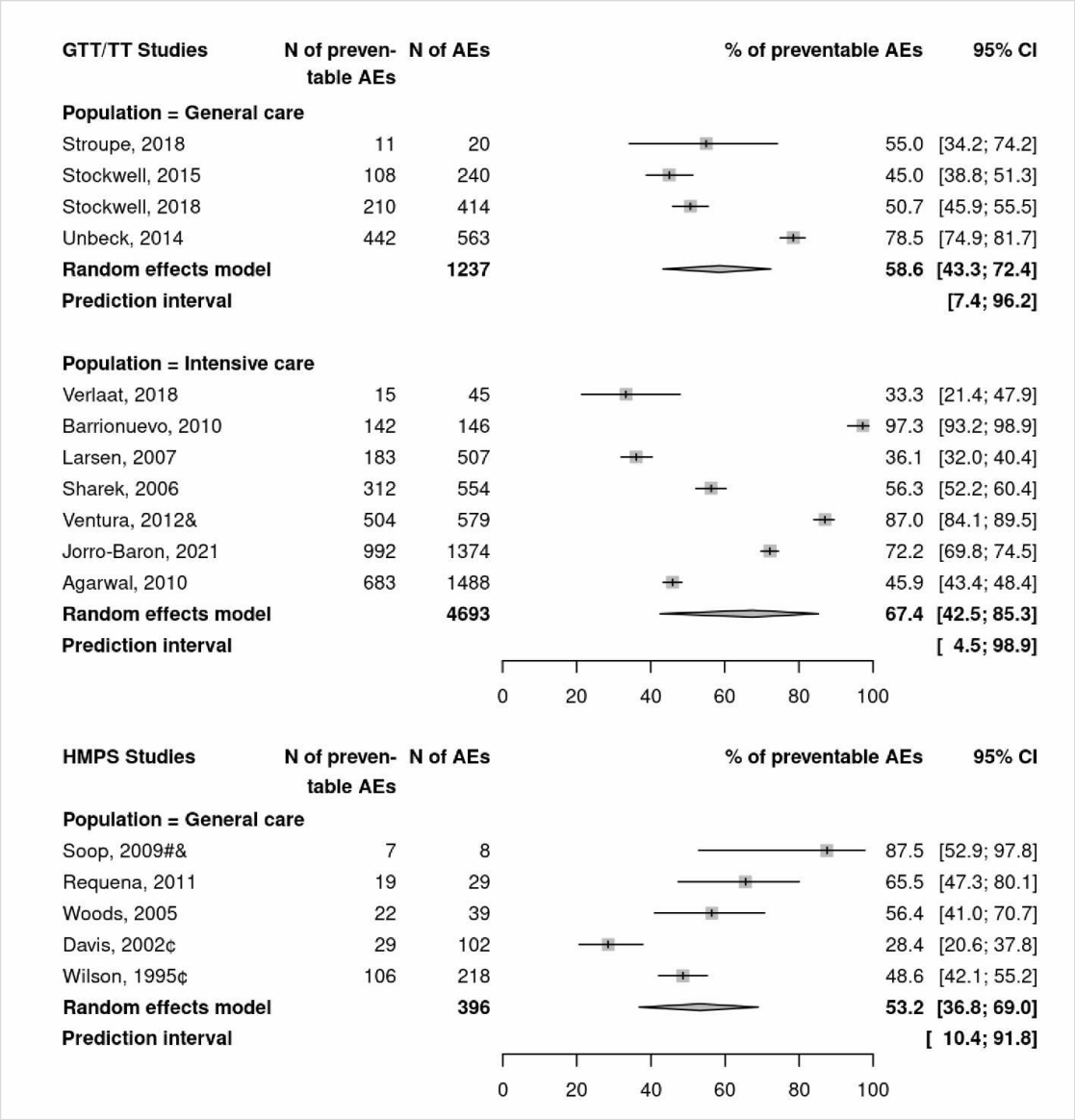


Figure S4 Forest plot of percentage of preventable adverse events (AEs) by general care respective intensive care population and methodology, ordered by number of AEs.[16-20 24-26 28 30-34 37 38 40 41 43-45]
Calculation of total number of AEs.
& Calculation of number of preventable AEs.
¢ Scored 2–6 on the causation scale compared to 4–6 for other studies using this scale to determine whether an AE was caused by healthcare management rather than the patient´s disease. However, the number of preventable AEs are the ones scored 4–6 on the preventability scale.
GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; TT, Trigger Tool.

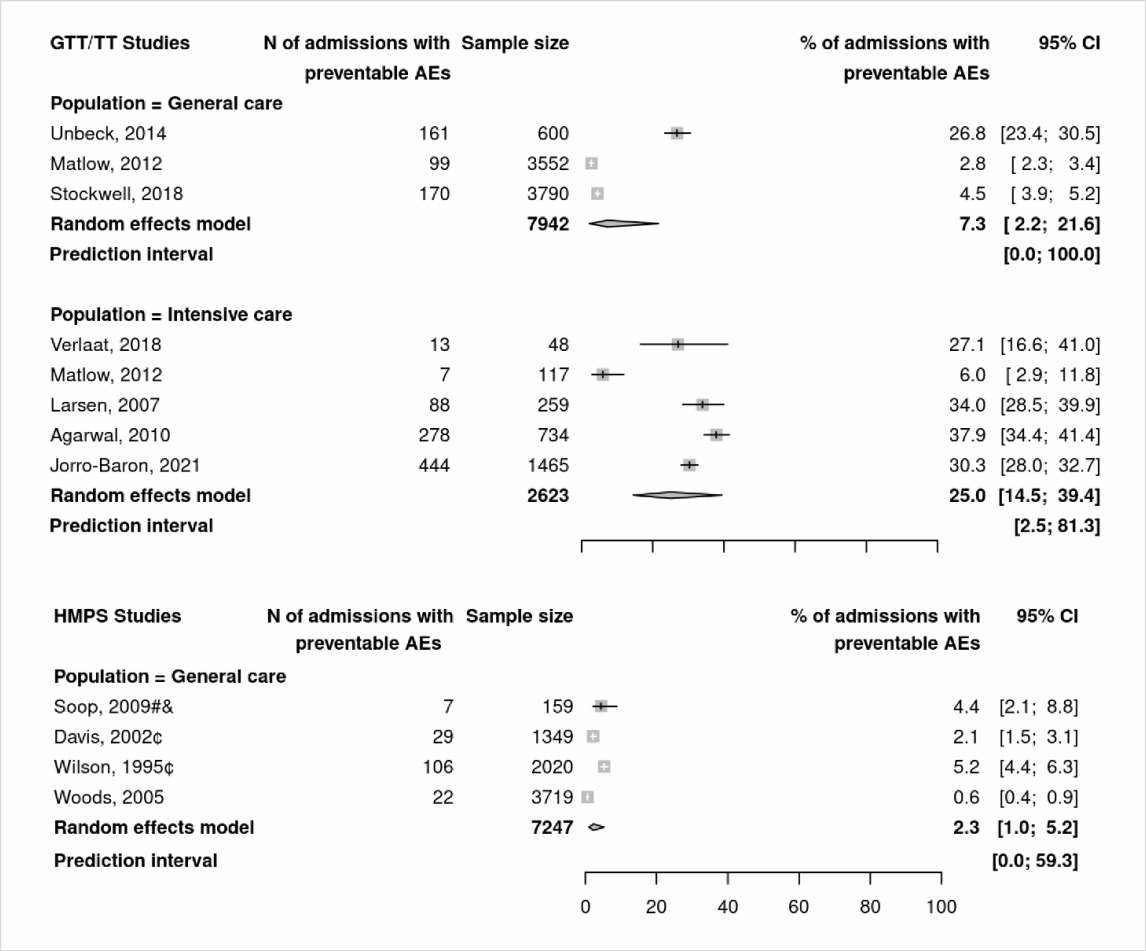


Figure S5 Forest plot of percentage of admissions with preventable adverse events (AEs) by general care respective intensive care population and methodology, ordered by sample size. [7-9 17 18 20 24-26 30-34 37 40 41 45]
Calculation of total number of AEs.
& Calculation of number of preventable AEs.
¢ Scored 2–6 on the causation scale compared to 4–6 for other studies using this scale to determine whether an AE was caused by healthcare management rather than the patient´s disease. However, the preventable AEs are the ones scored 4–6 on the preventability scale.
GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; TT, Trigger Tool.

Quality assessment tool, quality assessments and sensitivity analysis

Adaptation of the quality assessment tool

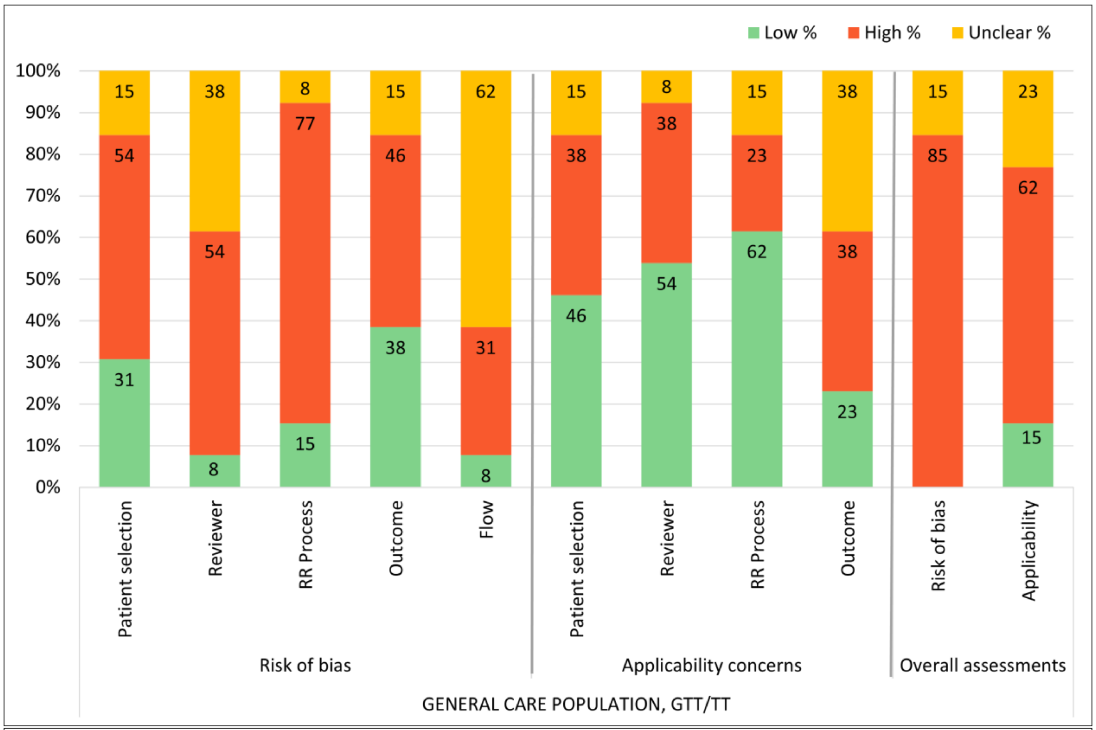
From the Quality Assessment Tool for Diagnostic Accuracy Studies 2 (QUADAS-2) we adopted the use of domains with domain specific signalling questions which provide domain specific judgements on the perceived risk of bias. We also used domain specific questions addressing concerns regarding applicability. The latter was used to judge whether the primary study was suitable for our review objective. We added a study level assessment of both the risk of bias and applicability concerns across all domains. The research team searched for empirical evidence from published systematic reviews addressing sources of bias and variation in studies evaluating record review methods for AE detection. We also discussed which additional sources could affect estimates of the incidence of AEs on theoretical grounds, and verified which domains and signalling questions of QUADAS-2, the tool described by Musy et al.[47] and later by Eggenschwiler et al.[48] were relevant to the current review. Based on these efforts, we created signalling questions and grouped them into domains. The tool was tested on a sample of primary studies nearly meeting our inclusion criteria and was then fine-tuned to enhance clarity and ease of use. The final tool consists of five domains: patient selection, reviewers, record review process, outcomes, and flow (supplementary material 2). Each domain includes two to three signalling questions which form the basis for the assessment of risk of bias and applicability-related concerns. These were rated as either low, high, or unclear.

The quality assessment ratings for each study are shown in table S7. The ratings per domain and overall ratings for the general population and all methodologies are shown in figure S6. Ratings per domain for the intensive care population is shown in figure S7.

Table S7 Assessment of risk of bias and concerns regarding applicability, for each of the five domains

First author, publication year	Patient selection		Reviewer		Record review process		Outcomes		Flow	Overall risk of bias	Overall Concerns regarding applicability
	Risk of bias	Applicability concerns	Risk of bias	Applicability concerns	Risk of bias	Applicability concerns	Risk of bias	Applicability concerns	Risk of bias		
General care population, GTT/TT methodology											
Chapman, 2014 [5]	High	Unclear	High	High	High	Low	Unclear	Low	Unclear	High	High
Davenport, 2017 [6]	High	High	High	High	High	Unclear	High	High	Low	High	High
Fajreldines, 2019 [10]	Low	Low	High	High	High	High	High	High	High	High	High
Kirkendall, 2012 [11]	Low	Low	Unclear	Low	Unclear	Low	Low	Low	Unclear	Unclear	Low
Matlow, 2012 [13-15]*	Low	Low	Low	Low	High	Low	Low	Unclear	Unclear	High	Unclear
Paredes Esteban, 2015 [16]	High	Low	High	Low	High	Low	High	Unclear	High	High	Unclear
Salimath, 2020 [18 19]	High	High	Unclear	High	High	High	High	Unclear	Unclear	High	High
Shah, 2009 [20 21]	High	Low	High	Low	Low	Low	High	Unclear	Unclear	High	Unclear
Solevag, 2014 [22]	High	High	High	High	High	High	Low	High	High	High	High
Stockwell, 2015 [25]	Unclear	High	Unclear	Low	High	Low	Unclear	Unclear	Unclear	High	High
Stockwell, 2018 [26 27]	High	Unclear	High	Unclear	High	Low	Low	High	Unclear	High	High
Stroupe, 2018 [28]	Unclear	High	Unclear	Low	High	Unclear	High	High	High	High	High
Unbeck, 2014 [29]	Low	Low	Unclear	Low	Low	Low	Low	Low	Unclear	Unclear	Low
General care population, HMPS methodology											
Brennan, 1991 [2-4]	Low	Low	Unclear	Unclear	Low	Low	Low	Unclear	Low	Unclear	Unclear
Davis, 2002 [7-9]	Low	Low	Unclear	Low	Low	Low	Low	Unclear	Low	Unclear	Unclear
Letaief, 2010 [12]	Low	Low	High	Unclear	High	Unclear	Unclear	Unclear	Low	High	High
Requena, 2011 [17]	High	Low	High	Low	High	Low	High	Unclear	High	High	Unclear
Sommella, 2014 [23]	Low	Low	Unclear	Low	Unclear	Low	Unclear	Unclear	Unclear	High	High
Soop, 2009 [24]	Low	Low	Unclear	Low	Low	Low	Low	Low	Low	Unclear	Low
Wilson, 1995 [30]	Low	Low	Unclear	Low	Low	Low	Low	Unclear	Low	Unclear	Unclear
Woods, 2005 [31-33]	Low	Low	Low	Low	High	Low	Low	Unclear	Unclear	High	Unclear
Zegers, 2009 [34 35]	Low	Low	Unclear	Low	Low	Low	Low	Low	Low	Unclear	Low
Intensive care population, GTT/TT methodology#											
Agarwal, 2010 [36]	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	Unclear	High	Unclear
Barrionuevo, 2010 [37]	Low	Low	Unclear	Low	High	Low	Unclear	Unclear	High	High	Unclear
Hooper, 2014 [38]	Low	Unclear	High	Low	High	Low	Low	Unclear	Unclear	High	Unclear
Jorro-Baron, 2021 [39]	Low	Low	Low	Low	High	Low	High	Unclear	Unclear	High	Unclear
Larsen, 2007 [40]	Low	Low	High	Low	High	Low	High	Unclear	Unclear	High	Unclear
Maziero, 2020 [41]	High	Unclear	High	Unclear	High	Unclear	High	Unclear	Unclear	High	High
Sharek, 2006 [42]	Low	Low	Low	Low	Unclear	Low	High	Low	Unclear	High	Low
Ventura, 2012 [43]	Low	High	High	Unclear	High	High	High	High	Low	High	High
Verlaat, 2018 [44]	High	High	Unclear	Unclear	High	Low	Low	Low	High	High	High
Vermeulen, 2014 [45]	Low	Low	Unclear	Low	High	Unclear	High	Low	High	High	Unclear

* Assessment based on overall outcome.
# No study using the HMPS methodology included mainly intensive care patients.
GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; TT, Trigger Tool.



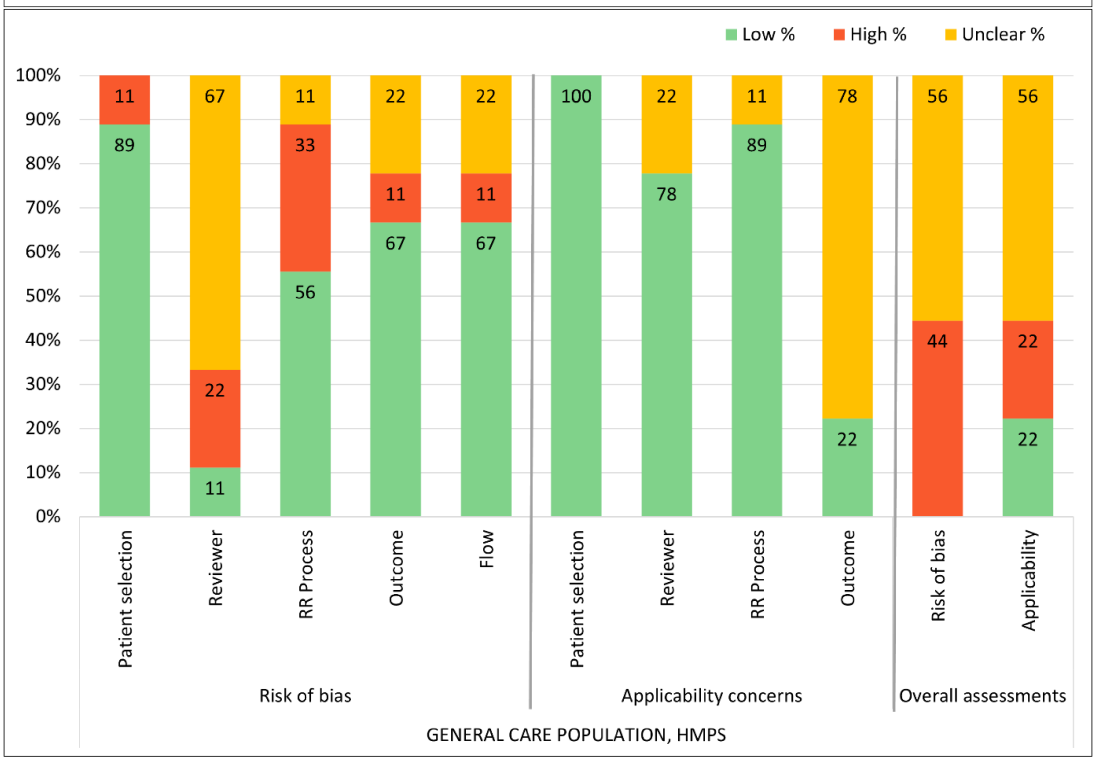


Figure S6 Quality assessment outcomes of included studies (n=22), for general care population stratified by methodology (GTT/TT n=13 respective HMPS n=9).
Presented in risk of bias and concerns regarding applicability for the five defined domains. Summarized in overall assessments by record review method.
Matlow et al.[8] was assessed for the overall cohort and included in the general care population.
GTT, Global Trigger Tool; HMPS, Harvard Medical Practice Study; RR, record review; TT, Trigger Tool.

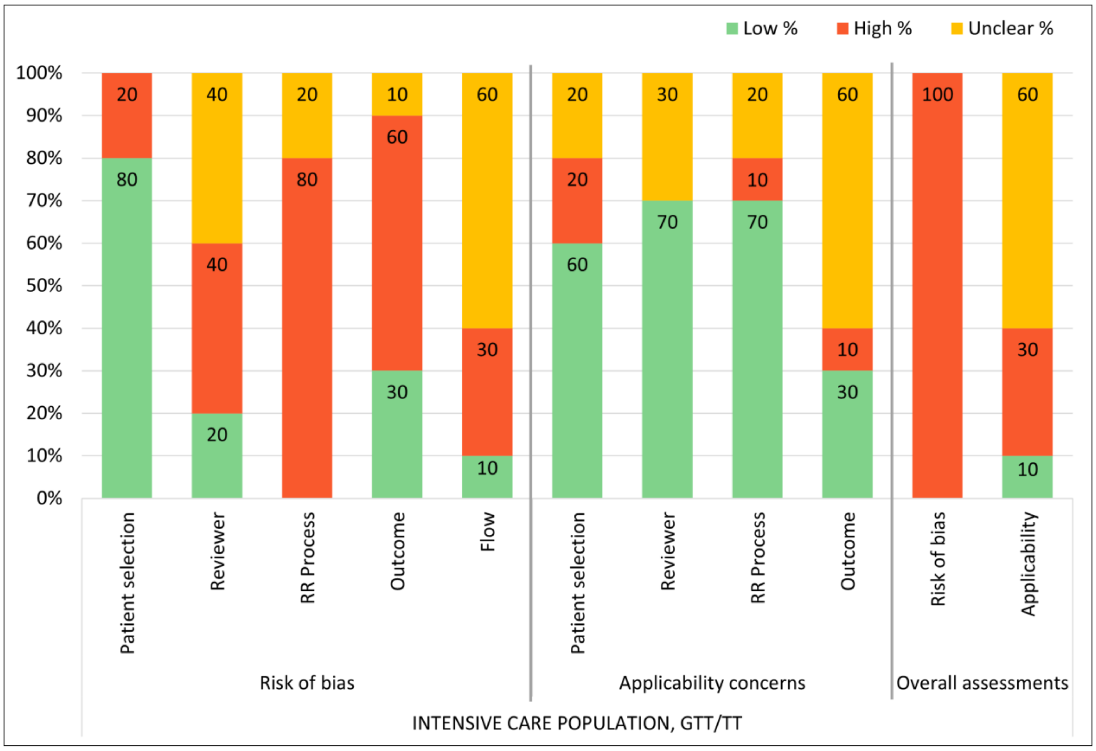


Figure S7 Quality assessment outcomes of included studies in intensive care (n=10) population (GTT/TT).
Presented in risk of bias and concerns regarding applicability for the five defined domains. Summarized in overall assessments for the two population subgroups.
Matlow et al.[8] was assessed for the overall cohort and included in the general care population.
GTT, Global Trigger Tool; RR, record review; TT, Trigger Tool.

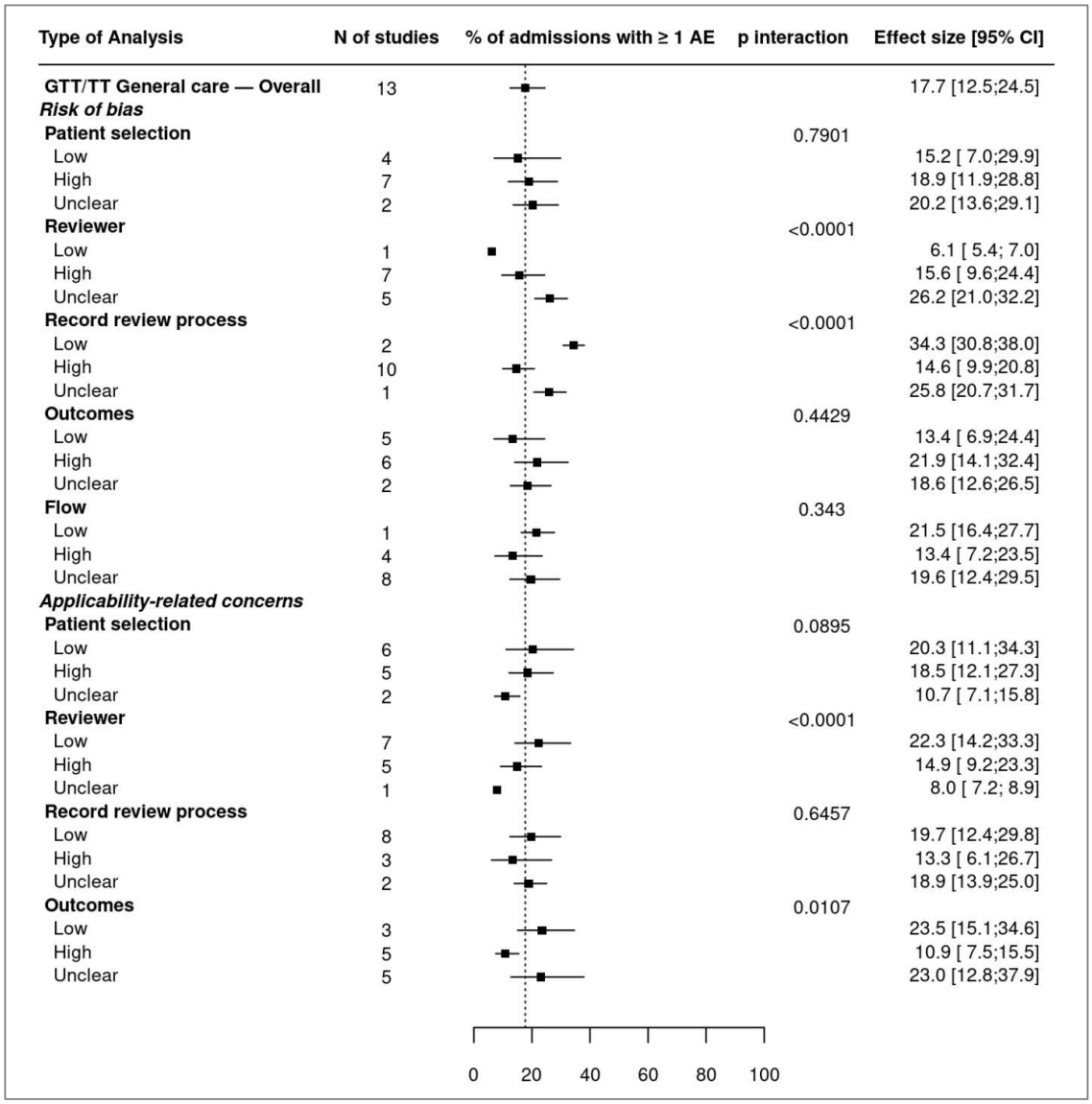


Figure S8 Forest plot with stratified analysis of the risk of bias and applicability-related concerns for general care population by GTT/TT methodology.
AE, adverse event; GTT, Global Trigger Tool; TT, Trigger Tool.

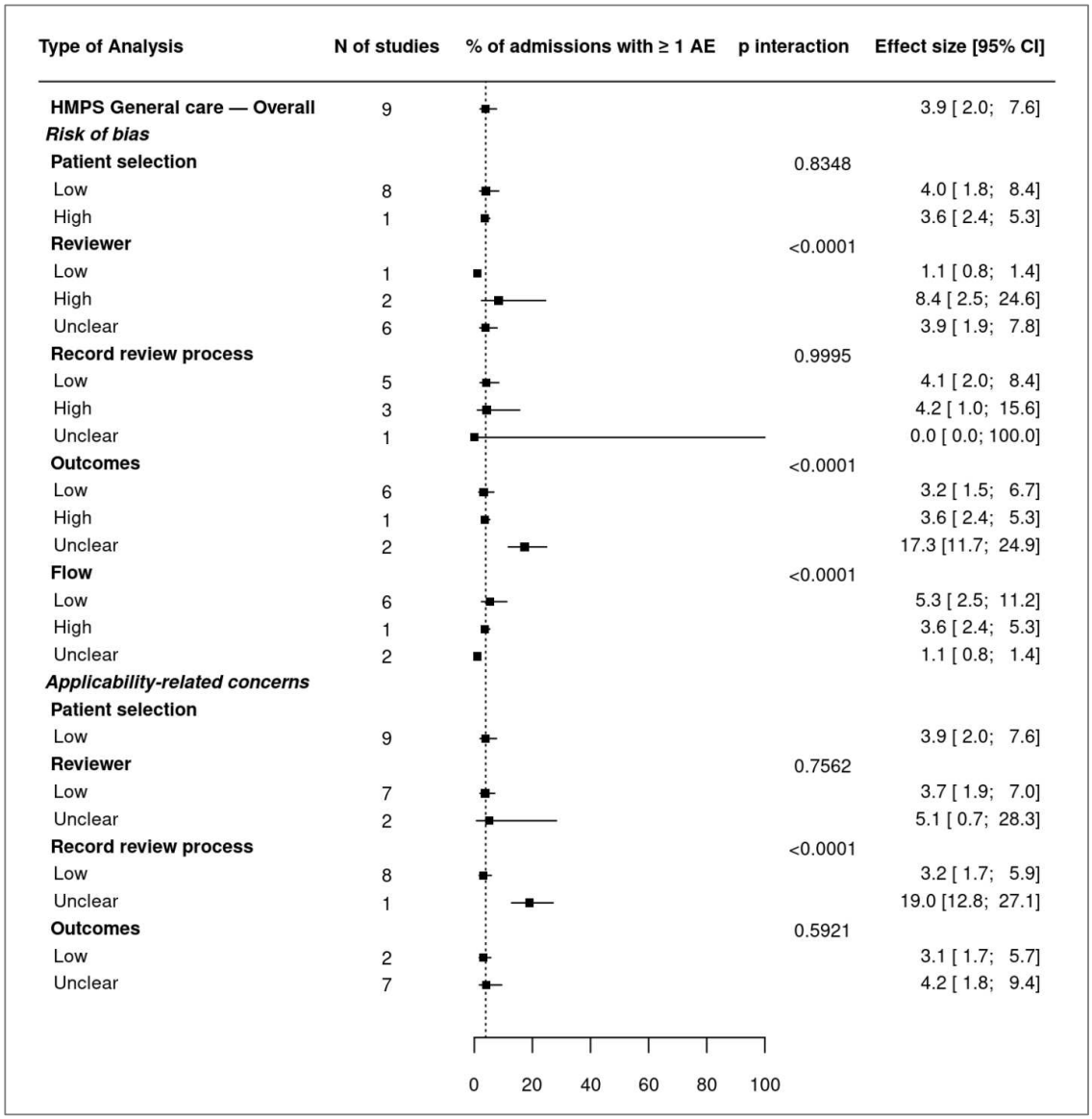


Figure S9 Forest plot with stratified analysis of the risk of bias and applicability-related concerns for general care population by HMPS methodology.
AE, adverse event; HMPS, Harvard Medical Practice Study.

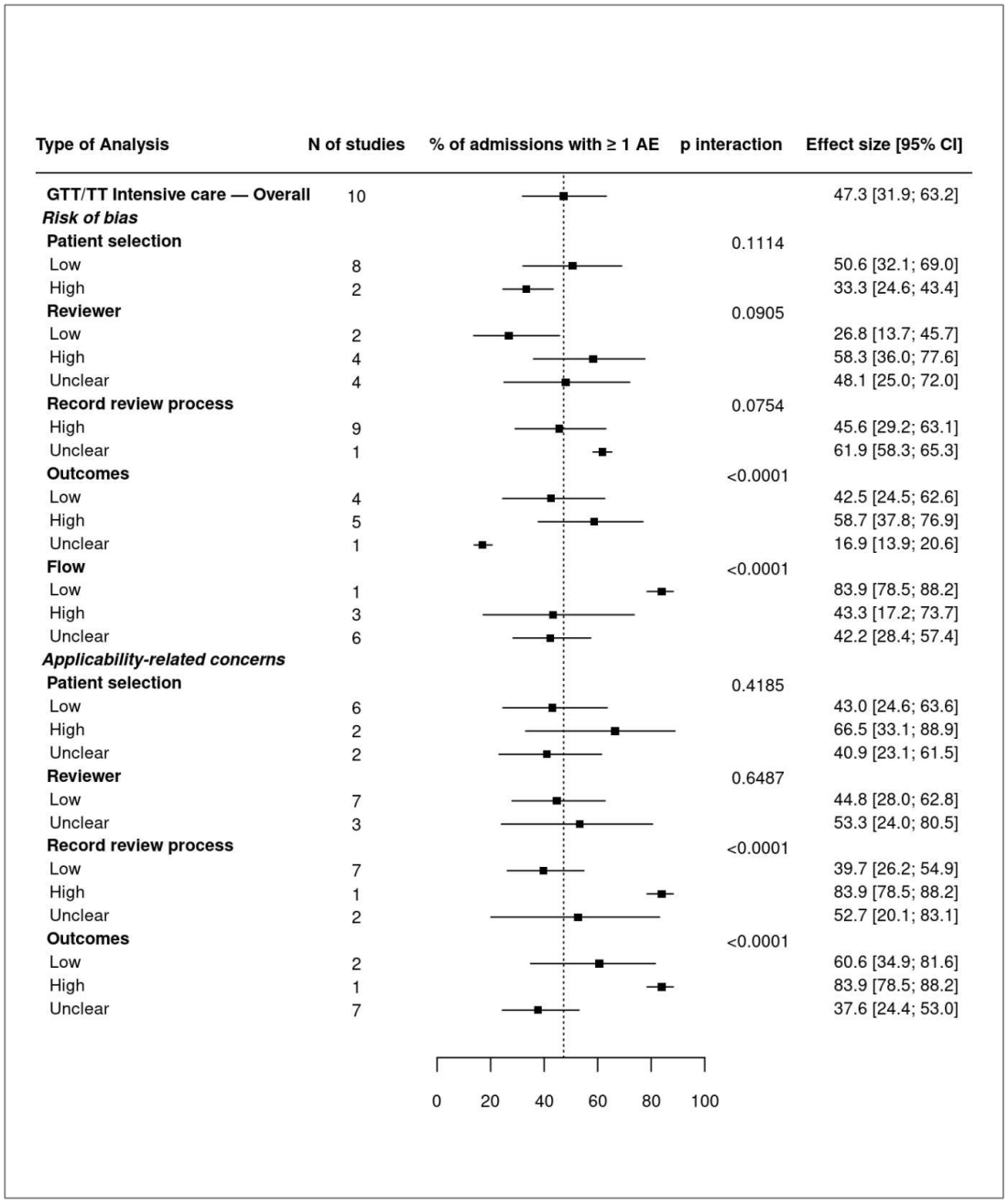


Figure S10 Forest plot with stratified analysis of the risk of bias and applicability-related concerns by intensive care population (GTT/TT).
AEs, adverse events; GTT, Global Trigger Tool; TT, Trigger Tool.

References

1 Bramer WM, Giustini D, de Jonge GB*, et al.* De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc* 2016;104:240-3.

2 Page MJ, McKenzie JE, Bossuyt PM*, et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.

3 Chapman SM, Fitzsimons J, Davey N*, et al.* Prevalence and severity of patient harm in a sample of UK-hospitalised children detected by the Paediatric Trigger Tool. *BMJ Open* 2014;4:e005066.

4 Davenport MC, Dominguez PA, Ferreira JP*, et al.* Measuring adverse events in pediatric inpatients with the Global Trigger Tool. *Arch Argent Pediatr* 2017;115:357-63.

5 Fajreldines A, Schnitzler E, Torres S*, et al.* Measurement of the incidence of care-associated adverse events at the Department of Pediatrics of a teaching hospital. *Arch Argent Pediatr* 2019;117:e106-e09.

6 Kirkendall ES, Kloppenborg E, Papp J*, et al.* Measuring adverse events and levels of harm in pediatric inpatients with the Global Trigger Tool. *Pediatrics* 2012;130:e1206-14.

7 Matlow A, Flintoft V, Orrbine E*, et al.* The development of the Canadian paediatric trigger tool for identifying potential adverse events. *Healthc Q* 2005;8:90-93.

8 Matlow AG, Baker GR, Flintoft V*, et al.* Adverse events among children in Canadian hospitals: the Canadian Paediatric Adverse Events Study. *CMAJ* 2012;184:E709-18.

9 Matlow AG, Cronin CM, Flintoft V*, et al.* Description of the development and validation of the Canadian Paediatric Trigger Tool. *BMJ Qual Saf* 2011;20:416-23.

10 Paredes Esteban RM, Garrido Perez JI, Ruiz Palomino A*, et al.* Implementation of a Plan of Patient Safety in Service of Pediatric Surgery. First results. *Cir Pediatr* 2015;28:111-17.

11 Geetanjali S, Ganachari MS. Development of a pediatric focused trigger tool to assess the prevalence of adverse events at a hospital setting: A retrospective structured chart review. *IJPR* 2020;12:1387-98.

12 Salimath GS, Ganachari MS, Gudhoor M. Paediatric Focused Triggering Tool (PFTT) To Assess the Harm and its Utilization to Minimize the Levels of Harm among Children at a Tertiary Care Hospital. *IJPER* 2020;54:819-25.

13 Lander L, Roberson DW, Plummer KM*, et al.* A trigger tool fails to identify serious errors and adverse events in pediatric otolaryngology. *OHNS* 2010;143:480-86.

14 Shah RK, Lander L, Forbes P*, et al.* Safety on an inpatient pediatric otolaryngology service: many small errors, few adverse events. *Laryngoscope* 2009;119:871-9.

15 Solevag AL, Nakstad B. Utility of a Paediatric Trigger Tool in a Norwegian department of paediatric and adolescent medicine. *BMJ Open* 2014;4:e005011.

16 Stockwell DC, Bisarya H, Classen DC*, et al.* A trigger tool to detect harm in pediatric inpatient settings. *Pediatrics* 2015;135:1036-42.

17 Landrigan CP, Stockwell D, Toomey SL*, et al.* Performance of the Global Assessment of Pediatric Patient Safety (GAPPS) Tool. *Pediatrics* 2016;137:06.

18 Stockwell DC, Landrigan CP, Toomey SL*, et al.* Adverse Events in Hospitalized Pediatric Patients. *Pediatrics* 2018;142:08.

19 Stroupe LM, Patra KP, Dai Z*, et al.* Measuring Harm in Hospitalized Children via a Trigger Tool. *J Pediatr Nurs* 2018;41:9-15.

20 Unbeck M, Lindemalm S, Nydert P*, et al.* Validation of triggers and development of a pediatric trigger tool to identify adverse events. *BMC Health Serv Res* 2014;14:655.

21 Brennan TA, Leape LL, Laird NM*, et al.* Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:370-76.

22 Hiatt HH, Barnes BA, Brennan TA*, et al.* A study of medical injury and medical malpractice. *N Engl J Med* 1989;321:480-4.

23 Leape LL, Brennan TA, Laird N*, et al.* The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med* 1991;324:377-84.

24 Davis P, Lay-Yee R, Briant R*, et al.* Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J* 2002;115:U271.

25 Davis P, Lay-Yee R, Briant R*, et al.* Adverse events in New Zealand public hospitals II: preventability and clinical context. *N Z Med J* 2003;116:U624.

26 Davis P, Lay-Yee R, Briant R*, et al.* Adverse events on New Zealand public hospitals: Principal findings from a national survey. Occasional Paper No 3: Wellington: Ministry of Health, 2001.

27 Letaief M, El Mhamdi S, El-Asady R*, et al.* Adverse events in a Tunisian hospital: results of a retrospective cohort study. *Int J Qual Health Care* 2010;22:380-85.

28 Requena J, Miralles JJ, Mollar J*, et al.* Clinical safety paediatric patients. *RCA* 2011;26:353-58.

29 Sommella L, de Waure C, Ferriero AM*, et al.* The incidence of adverse events in an Italian acute care hospital: findings of a two-stage method in a retrospective cohort study. *BMC Health Serv Res* 2014;14:358.

30 Soop M, Fryksmark U, Koster M*, et al.* The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. *Int J Qual Health Care* 2009;21:285-91.

31 Wilson RM, Runciman WB, Gibberd RW*, et al.* The Quality in Australian Health Care Study. *Med J Aust* 1995;163:458-71.

32 Thomas EJ, Orav EJ, Brennan TA. Hospital ownership and preventable adverse events. *Int J Health Serv* 2000;30:745-61.

33 Thomas EJ, Studdert DM, Burstin HR*, et al.* Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000;38:261-71.

34 Woods D, Thomas E, Holl J*, et al.* Adverse events and preventable adverse events in children. *Pediatrics* 2005;115:155-60.

35 Zegers M, de Bruijne MC, Wagner C*, et al.* Design of a retrospective patient record study on the occurrence of adverse events among patients in Dutch hospitals. *BMC Health Serv Res* 2007;7:27.

36 Zegers M, de Bruijne 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.

37 Agarwal S, Classen D, Larsen G*, et al.* Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Med* 2010;11:568-78.

38 Barrionuevo LS, Esandi ME. Epidemiology of adverse events in the neonatal unit of a regional public hospital in Argentina. *Arch Argent Pediatr* 2010;108:303-10.

39 Hooper AJ, Tibballs J. Comparison of a Trigger Tool and voluntary reporting to identify adverse events in a paediatric intensive care unit. *Anaesth Intensive Care* 2014;42:199-206.

40 Jorro-Baron F, Suarez-Anzorena I, Burgos-Pratx R*, et al.* Handoff improvement and adverse event reduction programme implementation in paediatric intensive care units in Argentina: a stepped-wedge trial. *BMJ Qual Saf* 2021;23:23.

41 Larsen GY, Donaldson AE, Parker HB*, et al.* Preventable harm occurring to critically ill children. *Pediatr Crit Care Med* 2007;8:331-36.
42 Maziero ECS, Cruz EDA, Alpendre FT*, et al.* Association between nursing work conditions and adverse events in neonatal and pediatric Intensive Care Units. *Rev Esc Enferm USP* 2020;54:e03623.
43 Sharek PJ, Horbar JD, Mason W*, et al.* Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics* 2006;118:1332-40.
44 Ventura CM, Alves JG, Meneses Jdo A. Adverse events in a Neonatal Intensive Care Unit. *Rev Bras Enferm* 2012;65:49-55.
45 Verlaat CW, van der Starre C, Hazelzet JA*, et al.* The occurrence of adverse events in low-risk non-survivors in pediatric intensive care patients: an exploratory study. *Eur J Pediatr* 2018;177:1351-58.
46 Vermeulen JM, van Dijk M, van der Starre C*, et al.* Patient safety in South Africa: PICU adverse event registration. *Pediatr Crit Care Med* 2014;15:464-70.
47 Musy SN, Ausserhofer D, Schwendimann R*, et al.* Trigger Tool-Based Automated Adverse Event Detection in Electronic Health Records: Systematic Review. *J Med Internet Res* 2018;20:e198.
48 Eggenschwiler LC, Rutjes AWS, Musy SN*, et al.* Variation in detected adverse events using trigger tools: A systematic review and meta-analysis. *PLoS One* 2022;17:e0273800.

Supplementary Material 2

Incidence and characteristics of adverse events in paediatric inpatient care: a systematic review and meta-analyses

Quality Assessment Tool for Global Trigger Tool, Trigger Tool or Harvard Medical Practice Study methodology studies

State your review question here and answer all signalling questions and judgements considering this review question.

Aim: The aim is to report incidence and characteristics of adverse events (AEs), in paediatric inpatient care, as detected with Global Trigger Tool (GTT), Trigger Tool or the Harvard Medical Practice Study methodology.

Tick the PICO that applies to the study report you aim to assess

□ PICO 1:

P Patient records to paediatric patients hospitalized to any paediatric ward with **normal level of inpatient care such as surgical, medicine or orthopaedic wards etc.**

I The manual retrospective record review (RRR) methods: GTT, a modified GTT version so called Trigger Tools (added/removed/modified triggers) or Harvard Medical Practice Study methodology.

C Not applicable.

O Incidence overall and by characteristics of AEs.

□ PICO 2:

P Patient records to paediatric patients hospitalized to any paediatric ward with **higher level of inpatient care such as paediatric intensive care units/neonatal intensive care units etc.**

I The manual retrospective record review (RRR) methods: GTT, a modified GTT version so called Trigger Tools (added/removed/modified triggers) or Harvard Medical Practice Study methodology.

C Not applicable.

O Incidence overall and by characteristics of AEs.

Dillner P, *et al. BMJ Qual Saf* 2022;0:1–17. doi: 10.1136/bmjqs-2022-015298

DOMAIN 1A: PATIENT SELECTION

Describe methods of patient selection:

Describe the rational for your judgement:

A. Risk of bias

Signaling questions:

1. ***Was a consecutive or random sample of patient records enrolled?***
   Yes / No / Unclear
   Reflect if all the subjects selected or recruited were from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants in a consecutive manner? If all accessible patient records were selected as sample or if the process of sampling was done with the method of random sampling, the question will be answered as "yes".

2. ***Are the selection criteria described in a way that the study population of interest is likely to be included?***
   Yes / No / Unclear
   This question will be answered with "no" when patients with different profiles are not considered. This typically occurs by exclusion at study entry. Such exclusions are likely to alter the estimates of incidence, potentially leading to over- or underestimations of the incidence of AEs. For example: exclusion of certain group of patients due to a long length of stay or with lots of transfers.

RoB judgement: Could the selection of patient records have introduced bias?
RISK: LOW / HIGH / UNCLEAR

B. Applicability:

If a study did not meet the patient population or setting as described in the review question, there will be a high concern regarding its applicability. In this specific review, we allow for a broad range of in-hospital specialities and care levels. If the case mix or types of specialties in the study do not fit the review question, this may result in a high concern of applicability. Only classify in *unclear* if the report does not specify the patient population.

Do not punish a study twice. If risk of bias is rated as high, applicability should be rated as high only if additional issues are raised.

Are there concerns that the included patients and setting do not match the review question?
CONCERN: LOW / HIGH / UNCLEAR

DOMAIN 1B: REVIEWER

State the determined number of reviewers:

Describe the reviewer characteristics (e.g., training, experience):

Describe the rational for your judgement:

A. Risk of bias
Signaling questions:

1. ***Were reviewers selected based on his/her relevant clinical background?***
   Yes / ***No*** / ***Unclear***
   The lack of experience of the reviewer(s) in the clinical setting may introduce bias. For reviewers with appropriate clinical background within pediatric care the bias might be lower.

2. ***Were reviewers trained on using trigger tool methodology and application?***
   Yes / ***No*** / ***Unclear***
   The lack of RRR training in applying the RRR may introduce bias. For reviewers with more training the bias might be lower.

3. **Did reviewers have experience in applying the targeted RRR methodology or another structured RRR methodology?**
   Yes / ***No*** / ***Unclear***
   The lack of RRR experience in applying the RRR may introduce bias. For reviewer with more RRR experience the bias might be lower.

RoB judgement: Could the selection of reviewers have introduced bias?
RISK: LOW / HIGH / UNCLEAR

B. Applicability*:*
For example, if the profile of reviewers applying the RRR method in the study substantially differ from the profile of health care professionals that would apply the RRR in clinical practice, a high concern may arise.

Do not punish a study twice. If risk of bias is rated as high, applicability should be rated as high only if additional issues are raised.

Are there concerns that the reviewers do not match the review question?
CONCERN: LOW / HIGH / UNCLEAR

DOMAIN 2: RECORD REVIEW PROCESS

Description of the record review process:

Describe the rational for your judgement:

A. Risk of Bias
Signaling questions:

1. ***Did the reviewers first judge/analyze the records independently (e.g., stage 1 screening) and then together?***
   Yes / No / Unclear
   Lack of independent review/screening in duplicate and consensus discussions may introduce bias. The risk of bias might be lower if the reviewers are working in pair with an independent review/screening stage and then a consensus stage.

2. **Did the reviewers have any kind of support during the review process?**
   Yes / No / Unclear
   Lack of support by a supervisor / senior during the review process may introduce bias. For reviewers with support the bias might be lower.

3. **Did the study have a monitoring/auditing process?**
   Yes / No / Unclear
   Lack of a monitoring/auditing process during the review process may introduce bias. The risk of bias in the review process might be lower if a monitoring/auditing process has been carried out.

RoB judgement: Could the conduct or interpretation of the trigger tool have introduced bias?
RISK: LOW / HIGH / UNCLEAR

B. Applicability*:*
If the RRR process or its implementation differ from your review question the results may not be applicable.

Do not punish a study twice. If risk of bias is rated as high, applicability should be rated as high only if additional issues are raised.

Are there concerns that the RRR process or its implementation differ from the review question?
CONCERN: LOW / HIGH / UNCLEAR

DOMAIN 3: OUTCOMES

Description of the definition of AE:

Describe how incidence was measured:

Describe the time frame of included AEs:

Describe the rational for your judgement, including if you deem definitions to be standard, or deviating from our review definitions:

A. Risk of Bias
Signaling questions:

1. ***Is the AE defined using a generally accepted definition?***
   Yes / No / Unclear
   IHI´s definition: "unintended physical injury resulting from or contributed by medical care that requires additional monitoring, treatment or hospitalization, or that results in death." (Griffin F, Resar R. IHI Global Trigger Tool for Measuring Adverse Events (second edition). Cambridge, Massachusetts: Instutute for Healthcare Improvment, 2009).

   HMPS´s definition: "unintended injury or complication which results in disability, death or prolonged hospital stay and is caused by health care management." (Wilson RM, Runciman WB, Gibberd RW, et al. The Quality in Australian Health Care Study. Medical Journal of Australia 1995;163(9):458-71).

2. ***Were AEs included if they occurred before, during and after index admission?***
   Yes / No / Unclear
   In the origin GTT and HMPS methodologies AEs are included following three inclusions periods (time frames). 1) the AE occurred before index admission and were the reason to index admission or were detect during index admission; 2) the AE occurred and were detected during index admission; and 3) the AE occurred during index admission and were detected after index admission (for example, an operation during index admission led to an infection or a pulmonary embolism postoperatively after discharge).
   Not following these inclusion periods may introduce bias. Such exclusions are likely to alter the estimates of incidence, potentially leading to underestimations of the incidence of AEs.

RoB judgement: Could the definition of outcomes have introduced bias?
RISK: LOW / HIGH / UNCLEAR

B. ***Applicability****:*
If the time frame of included AEs differs from your review question the results may not be applicable.

Do not punish a study twice. If risk of bias is rated as high, applicability should be rated as high only if additional issues are raised.

Are there concerns that the definition of outcomes differs from the review question?
CONCERN: LOW / HIGH / UNCLEAR

DOMAIN 4: FLOW

Description of the flow (number of patients included and number of patients analyzed):

Description of reasons to exclude patients from the statistical analyses:

Describe the rational for your judgement:

A. Risk of Bias

Signaling questions:

1. **Was the completeness of health records data adequate?**
   Yes / No / Unclear
   If yes, the risk of bias will be lower since it considers missing data for the analysis.

2. **Were all patients included in the analysis of incidence?**
   Yes / No / Unclear
   This question will be scored as "yes" if all patients who were recruited into the study were included in the analysis. No is scored if one or more patients are missing. High risk of bias may be judged if a substantial number of recruited patients is excluded from the statistical analysis. High risk of bias may be judged if reasons for exclusions of patients are reported and deemed likely to alter the estimates. If reasons are not reported, you may assume that one third of the missing would not be at random, and you should reflect if this may affect the study estimates importantly.

RoB judgement: Could the patient flow have introduced bias?
RISK: LOW / HIGH / UNCLEAR

Study level / Outcome level overall judgements

A. Risk of Bias

Guidance on how to reach your overall judgement	
Low risk of bias	• The study / result is judged to be at low risk of bias for all domains
High risk of bias	• If you have judged high risk of bias in one domain, it puts the study / result at overall high RoB OR • The study has a judgement of *UNCLEAR* in multiple domains in a way that substantially lowers confidence in the study / result
Some concerns	• The study has a judgement of *UNCLEAR* in at least one domain

Overall RoB judgement for the study:
RISK: LOW / HIGH / UNCLEAR

Tick one of the following options:

☐ **Study level judgement applies to all eligible outcomes**

☐ **Overall judgement depends on the outcome appraised**

If the overall judgement differs depending on the outcome assessed, describe the outcomes and each outcome specific overall judgements here:

Outcome	Overall judgement

B. Applicability

Guidance on how to reach your overall judgement	
Low concern	• The study / result is judged to be at low concern for all domains
High concern	• If you have judged high concern in one domain, it puts the study / result at overall high concern OR • The study has a judgement of *UNCLEAR* in multiple domains in a way that substantially lowers applicability of the study result(s)
Some concerns	• The study has a judgement of *UNCLEAR* in at least one domain

Overall judgement concerning applicability for the study
CONCERN: LOW / HIGH / UNCLEAR

Tick one of the following options:

☐ **Study level judgement applies to all eligible outcomes**

☐ **Overall judgement depends on the outcome appraised**

If the overall judgement differs depending on the outcome assessed, describe the outcomes and each outcome specific overall judgements here:

Outcome	Overall judgement