# Editorial

## Deaths by horsekick in the Prussian army – and other 'Never Events' in large organisations

In this issue of the journal, Moppett and Moppett analyse the distribution of 'Never Events' (i.e. serious adverse incidents like wrong site surgery) in the National Health Service (NHS) in the UK, and conclude that they follow the Poisson family of distributions [1]. This editorial discusses some pertinent questions related to this finding:

1 What does a 'Poisson distribution' imply?
2 When should we worry? (i.e. when, if ever, does a high rate or particular pattern of Never Events reflect fundamentally poor care in the organisation?)

### What does the Poisson distribution imply?

It was arguably the Prussian military that first recognised the notion of a 'never event' when it became concerned (in peacetime) about the number of its officers killed by horsekick in the cavalry [2, 3]. These were clearly serious incidents, at risk of lowering morale, potentially reflecting badly on the organisation and possibly a result of poor leadership or systems failure. Yet the Prussians had the foresight to collect hard data. Rather than simply blaming the victims for carelessness, the horses for viciousness, or the generals for incompetent leadership, the mathematician Ladislaus Bortkiewicz, a German of Polish descent, studied the death rates over a 20 year period (1875-1894). In a groundbreaking book entitled the '*Law of Small Numbers*' (a phrase that still resonates in statistical circles) [3], he showed that the data followed the 'Poisson distribution' (Fig. 1), the mathematics of which have been discussed several times previously in the pages of this journal [4–6].

Bortkiewicz went on to explain how this is a formal mathematical description of random, rare events. In other words, any real distribution (e.g. Prussian horsekicks) that resembles a Poisson in turn can be readily assumed to have arisen through chance events and not as a result of intent or design (e.g. inherent systems failure in the military). Similarly, if Moppett and Moppett have concluded that NHS Never Events are also Poisson-distributed and so random, then does it follow that they have no 'cause'

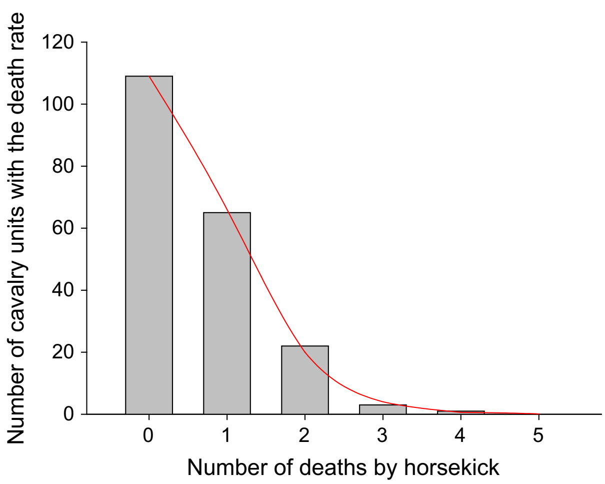


**Figure 1** Bortkiwewicz's data on deaths in Prussian cavalry. The bars represent the actual data with the y-axis simplified to show that in the majority of corps-years, there were no deaths; in 65 there was one death and so on. The red line is the predicted Poisson distribution, modelled on random rare events. The data were obtained by Bortkiewicz from the painstakingly collected *Pruessische Statistik* (Prussian Statistical Data), by the Imperial Statistical Bureau (available at: http://onlinebooks.library.upenn.edu/webbin/serial?id=prussstats).

This editorial accompanies the guidelines by the AAGBI Working Party, Anaesthesia, doi:10.1111/anae.13316.

 

and, in turn, that they cannot be prevented? We discuss below why this does not follow.

## *What does 'random' mean? Individual determinism vs collective probability*

A common Oxford entrance interview question asks candidates: "*What is the opposite of 'random'*?" An unhelpful (and likely unsuccessful) answer is simply to say 'non-random', but a more constructive approach might lead to a discussion of 'determinism'. That is, where all predisposing and causative factors to an event are known, its outcome is then predictable based on certain laws of nature. A candidate might reasonably offer the view that something that is deterministic cannot be truly random and they might even touch upon the notion of 'free will'.

However, the two concepts (randomness and determinism) are not as mutually exclusive as they first appear to be. The motion of one billiard ball (or gas molecule) when struck by another is an entirely deterministic event and no statistics are involved. If one knows exactly the mass, velocity, angle of incidence, friction, etc of the striking ball, the subsequent movement of the struck ball can always be calculated precisely.

If we tip a bag of balls onto the billiard table we could, given time and knowledge, calculate the movement and final position of each ball. However, using a different approach based on assumption of random behaviour, we can use statistical modelling to estimate the movement of all the balls at once and thereby predict, for example, the relative probabilities that they will end up being scattered away from the point at which they were tipped, versus remaining motionless or forming uniform patterns. This statistical approach requires no detailed knowledge of the forces applied to any single ball. Correspondingly, for example, in physics the temperature or pressure of a gas is a comment on the average velocity of the molecules within it, and not a description of the velocity of any given molecule. If we heat a gas, we cannot know if the velocity of any given molecule will increase; but we can be sure that the average velocity will rise. This mathematical approach is remarkably accurate (as the Nobel laureate physicist Eugene Wigner termed 'unreasonably so') despite its assumptions of randomness, in describing otherwise deterministic physical and biological behaviours [7].

In this sense, the concept of 'random behaviour' arises because we do not possess all the information about each discrete unit we are observing. Nevertheless, we are able to describe very accurately their behaviour using probabilistic modelling applied to all the units collectively. This statistical/probabilistic approach does not undermine – and is entirely compatible with - the fact that the behaviour of each individual unit could be decided deterministically. As humans, we generally choose our partners very carefully; yet the distribution of our alleles is best modelled as if we were mating randomly [8]. In this way, randomness (an assumption at the heart of the probabilistic approach) and determinism (the forces that drive an individual event) can co-exist in harmony.

Thus, one description of 'randomness' is where the collective behaviour of units in a population can be modelled statistically, based on an assumption that each of two or more possible outcomes is equally likely. The 'opposite of random' is a pattern of behaviour which is not compatible (i.e. statistically improbable) with this assumption. A concrete example of this distinction is the recent analysis by Carlisle, who concluded that the data of Fujii were very unlikely to have been derived from random sampling, and were therefore more likely deterministically derived [9–12].

## When should we worry?

So, if a Poisson distribution implies that the data follow a random distribution for rare events, we can now turn to our next question: when should we worry?

The answer depends on our perspective. If we are the patient (or the family), then none of the statistics matter: the event has happened to us and naturally we are always concerned. Therefore, the hospital quoted in Moppett and Moppett's paper that classed 11 wrong implants in ophthalmic surgery as one Never Event was wrong to do so. Similarly if we are the medical team, then the overall statistics offer little consolation. Just because Never Events across the country are randomly distributed, it does not mean (as discussed above) our own Event did not have an identifiable or preventable set of predisposing factors. It is imperative that we analyse these (root cause analysis, etc) and learn from them [13, 14].

If, however, we are health service planners or regulators, then our perspective may differ. We may be more concerned with identifying outlier hospitals or teams with inherent organisational problems (e.g. identifying a 'Mid-Staffs' [15]). Where the overall data follow a random (Poisson) distribution, this is notoriously difficult to do using statistics alone. The reasons for this are complex. For example, reviewers' comments on Moppett and Moppett's paper included discussions around Markov chain analysis, Firth-type bias-reduced logistic regression, relogit package in Zelig and the 'synthetic minority oversampling technique'. Readers will be relieved to hear we will not rehearse those arguments. Different analyses can yield opposite results, where hospitals identified as poor performers by one method are statistically no different from a hospital marked as a high performer. Problems of interpretation can arise where sample size is inevitably low (as it is with Never Events) and confidence intervals can be wide [16–18].

One additional confounder is that the human mind (even an expert one) is tempted to identify clusters as signifying non-random events. Recently, Sir Roger Boyle, Director of the National Institute for Cardiovascular Outcomes Research (and 'heart tsar') sent the Medical Director of the NHS some data that indicated a cluster of paediatric cardiac surgery deaths in Leeds, as a result of which operations there were suspended as 'unsafe'. A later re-analysis (that included the use of Poisson modelling) indicated there was no reason to suppose this was anything other than random variation, and surgery re-commenced (see: http://www.england.nhs.uk/2013/04/12/reports-chs-leeds/). Counter-intuitively, clusters are therefore often reassuring that random processes have occurred. Because we know that tossing a fair coin results, in the long run, in an equal number of heads and tails, it is often erroneously concluded from this that 'randomness' means an ordered or equal distribution of outcomes. Figure 2 might represent some hypothetical final positions of billiard balls spilled onto a table, or the geographical distribution of Never Events. It is Panel B, not Panel A, that should reassure us that random (Poisson) processes are at play. Panel A (a uniform distribution) suggests some planning or collusion. So one trigger for worry is if we see unusual, non-Poisson distributions (eg, uniform or bimodal) as these could indicate underlying problems in the system as a whole.

Another cause of worry is the absolute rate of Never Events. The Poisson distribution applies to rare, random events. However, this means that it applies equally well to Prussian officers kicked to death by horses as it does to the much less common event of horses kicked to death by their Prussian officers. The distribution of cases where the patient is accidentally awake during surgery is Poisson distributed across hospitals [19,

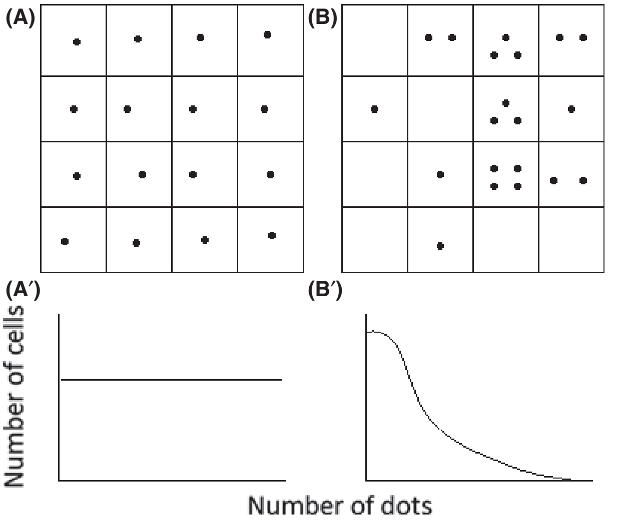


**Figure 2** Illustration of the appearance of random vs non-random distributions (a grid has been added for clarity). Panel A: the distribution of dots (billiard balls on a table or Never Events across a region) is even, described by the uniform distribution pattern in A′. Panel B: the dots show clustering in some cells, consistent with the Poisson distribution in B′. Note the similarity of B′ distribution with Figure 1. Thus perhaps counter-intuitively it is B/B′ that indicates random distribution, not A/A′ (See Supporting Information File S1 available online).

20]. So, probably, too are the cases where the anaesthetist is accidentally asleep [21]. The Poisson does not help us identify an acceptable rate of event; it only tells us that all these events arise randomly. To help a little with this problem, Weaver introduced the notion of the 'surprise index' (see Appendix), that mathematically separates the concept of 'rarity' from the surprise associated with that event arising [22]. Not all rare events are surprising: for example, the winner is always surprised on winning the lottery, but it is not a surprise that someone wins the lottery. On the other hand, it is both rare and a surprise to hear that a Prussian officer has kicked his horse to death, or that an anaesthetist has fallen asleep during surgery.

When the prevailing event rate increases, the shape of distribution changes gradually from Poisson to resemble a normal (Gaussian) distribution, so this pattern might be a trigger for concern: ie, a rare event has become a common event (Fig. 3).

## Summary

In summary, the interpretation of Moppett and Moppett's paper depends upon one's perspective (patient, physician or regulator) and upon one's understanding of randomness, determinism and probability. Moppett and Moppett have nicely shown that the distribution of Never Events in the NHS – just like that of deaths by horsekick in the Prussian cavalry over a century ago – follows a quasi-Poisson distribution. This means that the events, when viewed in totality, are rare and random. There is no reason to suspect any fundamental systems failure across the NHS or in any one hospital based on the Never Event data alone.

Yet, individual events can have specific causes. From an individual perspective we are reminded that, although we exercise our free will and determinism, we are surrounded by factors that can be regarded as random. The purpose of safe practice, embedded in professional guidelines, is to limit the translation of those factors into adverse outcomes and to mitigate any impacts. Thus, the number of car crashes on any given road may well be Poisson-distributed but this should not stop us from wearing seat belts, sticking to the speed limit, or investigating an accident. Accidental awareness may be described by a Poisson model, but this does not mean we stop checking the anaesthetic machine, monitoring end-tidal agent levels, or applying the now-recommended investigation and care pathways if it arises [23]. Similarly, Moppett and Moppett may have shown that the distribution of Never Events is random but this should not stop us from applying the World Health Organisation checklists, or from investigating by root cause analysis when an Event happens.

In the light of these statistical considerations, it is clear that Never Events are mis-named. They are really events that will *always* occur, and do so randomly - but from which learning for individuals and teams should emanate. Re-naming them Poisson-, or ideally Bortkiewicz-Events would better encapsulate these ideas, but it is perhaps rather too hopeful to think that this will catch on.

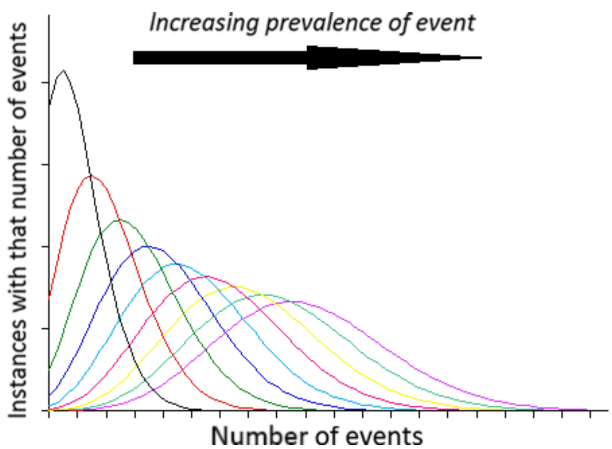


**Figure 3** Histograms of distribution by number of events, for an increasing prevalence of events (arbitrary scales, to show the general pattern). As the prevalence changes from rare (black line, extreme left distribution) to common (purple line, extreme right distribution), the pattern changes from Poisson to Gaussian.

## Conflicts of interest

Note that 'Bortkiewicz' has several spellings in the literature concerning his work. This is the Polish spelling, but his original work (reference 3) was published in the German version. Other versions exist. No external funding and no competing interests declared.

 

**J. J. Pandit**

*Consultant Anaesthetist/Professor & Fellow*
*Nuffield Department of Anaesthetics*
*Oxford University Hospitals NHS Trust, UK*
*St John's College*
*Oxford*

*Email: jaideep.pandit@dpag.ox.ac.uk*

## References

1. Moppett IK, Moppett SH. Surgical caseload and the risk of surgical Never Events in England. *Anaesthesia* 2015; **70**: doi: 10.1111/anae.13290.
2. Quine MP, Seneta E. Bortkiewicz's data and the Law of Small Numbers. *International Statistical Review* 1987; **55**: 173–81.
3. Bortkewitsch L. *Das Gesetz der Kleinen Zahlen*. Leipzig: G. Teubner, 1898.
4. Pandit JJ. Editorial: Gambling with ethics? A statistical note on the Poisson (binomial) distribution. *Anaesthesia* 2008; **63**: 1163–6.
5. Pandit JJ. If it hasn't failed, does it work? On the 'worst we can expect' from observational trial results, with reference to airway management devices. *Anaesthesia* 2012; **67**: 578–83.
6. Pandit JJ, Picton P, Mashour G. Penehcyclidine and awareness during anaesthesia: caution with zero numerators. *Anaesthesia* 2013; **68**: 121–35.
7. Wigner EP. The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics* 1960; **13**: 1–14.
8. Pandit JJ. Editorial: 'Hardy's Law' and genomics in Anaesthesia. *Anaesthesia* 2008; **63**: 1284–7.
9. Carlisle JB. The analysis of 168 randomised controlled trials to test data integrity. *Anaesthesia* 2012; **67**: 521–37.
10. Carlisle JB, Dexter F, Pandit JJ, Shafer SL, Yentis SM. Calculating the probability of random sampling for continuous variables in submitted or published randomised controlled trials. *Anaesthesia* 2015; **70**: 848–58.
11. Miller DR. Probability screening in manuscripts submitted to biomedical journals - an effective tool or a statistical quagmire? *Anaesthesia* 2015; **70**: 765–8.
12. Pandit JJ. On statistical methods to test if sampling in trials is genuinely random. *Anaesthesia* 2012; **67**: 456–462.
13. Harrop-Griffiths W. Never Events. *Anaesthesia* 2011; **66**: 158–62.
14. Gupta KJ, Cook TM. Accidental hypoglycaemia caused by an arterial flush drug error: a case report and contributory causes analysis. *Anaesthesia* 2013; **68**: 1179–87.
15. Department of Health. *Hard Truths: The Journey to Putting Patients First*. London: The Stationary Office, 2013.
16. Lilford R, Pronovost P. Using hospital mortality rates to judge hospital performance: A bad idea that just won't go away. *British Medical Journal* 2010; **340**: c2016.
17. Shahian DM, Normand ST. What is a performance outlier? *BMJ Quality and Safety* 2015; **24**: 95–9.
18. Normand ST, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: issues and applications. *Journal of the American Statistical Association* 1997; **92**: 803–14.
19. Pandit JJ, Cook TM, Jonker WR, O'Sullivan E. A national survey of anaesthetists (NAP5 Baseline) to estimate an annual incidence of accidental awareness during general anaesthesia in the UK. *Anaesthesia* 2013; **68**: 343–53.
20. Pandit JJ, Andrade J, Bogod DG, et al. 5th National Audit Project (NAP5) on accidental awareness during general anaesthesia: summary of main findings and risk factors. *Anaesthesia* 2014; **69**: 1089–101.
21. Robson R. Anaesthetist slept in chair as patient lay on operating table giving birth during major surgery. *Daily Mail* 21 March 2013 At: http://www.dailymail.co.uk/news/article-2296994/Anaesthetist-slept-chair-patient-lay-operating-table-giving-birth-major-surgery.html (accessed 12/08/2015).
22. Weaver W. Probability, rarity, interest, and surprise. *Scientific Monthly* 1948; **67**: 390–2.
23. Cook TM, Andrade J, Bogod DG, et al. 5th National Audit Project (NAP5) on accidental awareness during general anaesthesia: patient experiences, human factors, sedation, consent, and medicolegal issues. *Anaesthesia* 2014; **69**: 1102–16.
24. Hunter A, Parsons SD. *Applications of Uncertainty Formalisms*. Berlin: Springer, 2003.

## Appendix
## An example of Weaver's surprise index

Surprise is expressed by Weaver as [24]:

$$\text{Surprise} = \frac{\sum_i (P_i)^2}{P_r}$$

where the summation (Pi) is over all the possible outcomes and Pr is the probability that the given rare event actually occurs. A value of between 0 and 1 corresponds to a likely, unsurprising outcome; the larger the surprise value $> 1$, the greater the surprise.

In a lottery where *n* tickets are bought by the same number of people, the probability that any one person wins is 1/n. The surprise that someone, anyone, unknown to us wins the lottery is estimated by:

$$\text{Surprise} = \frac{n(1/n)^2}{(1/n)}$$

Which is 1; ie, this is not all that surprising. However, the probability that a named person will win the lottery (perhaps oneself) is the sum of them winning (1/n) and of them not winning $(1 - 1/n)$ as entered into the equation:

$$\text{Surprise} = \frac{(1/n)^2 + (1 - 1/n)^2}{1/n}$$

If only the named person buys the single ticket, $n = 1$ and the surprise is small. If n is large, the surprise is very large.

doi:10.1111/anae.13261

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**File S1.** More information about Figure 2. [Correction: File S1 was added after online publication on 23 November 2015, to provide more information]

Anaesthesia 2015 doi:10.1111/anae.13290

Original Article

# Surgical caseload and the risk of surgical Never Events in England

**I. K. Moppett[1,2] and S. H. Moppett[3]**

*1 Associate Professor and Honorary Consultant, Anaesthesia and Critical Care Section, Division of Clinical Neuroscience, Queen's Medical Centre, University of Nottingham, Nottingham, UK*
*2 Consultant, Department of Anaesthesia, 3 Divisional Nurse, Clinical Support Division, Nottingham University Hospitals, Nottingham, UK*

## Summary

Never Events are medical errors that are believed to be preventable with appropriate measures. We surveyed all English acute NHS trusts to determine the number of surgical Never Events and surgical caseload for 2011–2014. There were 742 surgically related Never Events in three years, with no change in the number annually. The risk of a surgical Never Event was 1 in 16 423 operations (95% CI 1 in 15 283 to 1 in 17 648) or 1 Never Event per 12.9 operating theatres per year (95% CI 1 in 12.1 to 1 in 13.9). The risk of severe harm due to a Never Event was approximately 1 in 238 939 operations. There was no meaningful association between number of Never Events and other safety indicators. Surgical Never Events are undoubtedly important to individual patients, but they are not a useful metric to judge quality of care.

*Correspondence to: I. K. Moppett*
*Email: iain.moppett@nottingham.ac.uk*
*Accepted: 21 September 2015*
*This original article is accompanied by an editorial by Pandit, Anaesthesia, doi: 10.1111/anae.13261.*

## Introduction

The UK National Patient Safety Agency (NPSA) introduced 'Never Events' into National Health Service practice in 2009. Analogous to 'non-pay' events in the USA, Never Events were defined as 'serious, largely preventable patient safety incidents that should not occur if the available preventative measures have been implemented' [1, 2]. Of the original eight categories of Never Events in the UK, 'wrong site surgery' and 'retained instrument postoperation' related directly to operating theatre practice. A third category of 'wrong implant' was introduced in 2011, and 'retained instrument' was redefined. The current list of surgical Never Events therefore consists of 'wrong site surgery', 'wrong implant/prosthesis' and 'retained foreign object postoperation' [3]. According to the NPSA a Never Event must, among others, fulfil the following criteria: there is existing national guidance and/or national safety recommendations on how the event can be prevented and support for implementation, and the event is largely preventable if the guidance is implemented [3].

By definition, there is no acceptable threshold for Never Events. The implication of national reports and guidance is that Never Events are sentinel events for poor care. There has been a significant political and media focus on Never Events in the UK, mirrored by experience from the USA. Organisations that experience Never Events may experience punitive action through financial consequences, increased scrutiny from regulatory authorities [4] and adverse publicity [5], and individuals involved may have disciplinary procedures invoked. Most importantly, patients may come to harm. The assumption of politicians [3] and commentators [6] appears to be that a greater number of Never Events is associated with poor practice at an organisation. It is intuitive that larger organisations may be more likely to have a Never Event, but this has never been investigated formally. This may be because denominator data are difficult to obtain; a recent report from NHS England stated that it was 'not possible to retrieve the total number (of patients) admitted to hospital for surgical care' [7].

An alternative approach to Never Events is to view them as random rare events; at a large scale such as NHS England they should occur relatively predictably, but at the small scale of the individual hospital they are unpredictable. We wished to investigate the occurrence of Never Events as random rare events, to test whether this approach would adequately model current data, and to provide testable statistical predictions based on the model.

## Methods

All English NHS acute hospital trusts providing surgical services were included. English acute care is provided by 'Trusts', which are the legal and managerial entity; individual Trusts may consist of one or more separate general or specialist hospitals. Never Events data are typically held and reported at Trust rather than hospital level. Data on Never Events were requested using the UK Freedom of Information Act, which places a legal obligation on public bodies to provide information when requested. Information about the number of surgical Never Events, type and severity of consequent harm [8] was requested from each Trust for two consecutive years (April 2011–March 2012 and April 2012–March 2013). Data for 2013–2014 were obtained from published reports from NHS England [9]. Data were generally only available at year level; only a few trusts provided data at more detailed level (quarterly or monthly). Two estimates of surgical caseload were also requested, the number of operations recorded in individual trust operating theatre management systems and the number of operating theatres (operating rooms). The data obtained using the Freedom of Information Act were cross-checked with data collated from other sources by NHS England, and publically available reports from NHS Trusts, healthcare commissioning bodies and the former Strategic Health Authorities. Where there was a discrepancy between the published data and the data provided by the Trust, this was clarified directly with the trust.

There is no single gold standard index for safety of hospital care [10]. We collated data for the Summary Hospital-level Mortality Index (SHMI) a UK-specific ratio of observed to expected (risk adjusted) in-hospital deaths and deaths within 30 day of hospital discharge [11, 12]. We also used the following 'safety metrics' from NHS Choices [13]: Care Quality Commission (CQC) rating (2014–2015) (1 – outstanding to 4 – inadequate); 'recommended by staff'; 'safe staffing' (ratio of actual to planned staffing); 'open and honest reporting' (a composite indicator of high incident reporting and staff perceptions of incident handling).

To provide a historical and geographical context, we reviewed the published literature for reported risks of Never Events using a combination of Medline searching, hand searching of reference lists, and forward searching of citations (Web of Science™, Google Scholar™ and publishers' websites).

### *Data analysis*

Our null hypotheses were:

$H_0$-1: number of Never Events is not predicted by organisation size (number of operations or number of operating theatres).

$H_0$-2: number of Never Events is not predicted by an index of overall patient outcome (SHMI).

$H_0$-3: number of Never Events in a year is not predicted by the safety metrics recorded by NHS Choices / CQC.

$H_0$-4: number of Never Events in a year is not predicted by the occurrence of Never Events previously.

All modelling was undertaken using the R statistics package [14]. We assumed that Never Events behave as rare random events linked to organisational size, and, as such, appropriately modelled using the Poisson/negative binomial family of distributions. The hypotheses were tested using the following approaches:

 

$H_0$-1-3: generalised linear modelling using Poisson and negative binomial models (`glm` and `glm.nb`) performed with the following predictor variables: number of operative procedures, number of theatres, SHMI and the NHS Choices safety metrics. The outcome was number of Never Events for the three years combined, and for each year separately. The association between number of operative procedures, number of operating theatres and SHMI was also assessed using simple linear regression. Sensitivity analysis was undertaken using alternative modelling approaches including adjusted Poisson inference estimates (sandwich), quasipoisson, hurdle and zero-inflated negative binomial models [15]; brief details of these alternative models are given in Appendix 1.

$H_0$-4: simple linear regression (`lm`) of number of Never Events in each year against the other two years.

As these were nested models, the Akaike Information Criterion (AIC) [16] was used to assess model suitability. The AIC is a statistical technique that attempts to find the optimal balance between model simplicity (number of predictors) and predictive ability.

Descriptive statistics were calculated for the number of recorded operations, operating theatres, Never Events, and associated harm. Estimate of Never Event risk (Never Events/number of operations and Never Events/number of operating theatres/year) were calculated using Wilson's exact method [17] to provide 95% CI. Two tailed 95% CI of the number of Never Events which would be expected from the modelled mean (lambda) of the regression models were calculated for each trust (`qpois`).

## Results

All 158 Trusts responded to the Freedom of Information Act request. Trusts that had merged or split during the period were able to provide data for individual hospitals. SHMI data were available for 140 trusts (excluding specialist and children's hospitals).

One hundred and thirty-eight Trusts were able to provide complete and reliable data for the number of operations performed; 20 Trusts were unable to provide complete data owing to a lack of or incomplete data for 2011–2012 (nine Trusts) or 2012–2013 (five Trusts), or were unable to separate out non-operating theatre-based procedures (nine Trusts). There were a median (IQR [range]) of 24 018 (15 591–31 520 [4007–88 160]) operations performed per Trust in 2011–2012 and 23 660 (14 938-30 111 [4055–88 160]) in 2012–2013. There were a total of 3200 operating theatres, with a median (IQR [range]) of 17 (12–26 [3–58]) per trust.

There were total of 742 surgically related Never Events reported by Trusts, with 255 in 2011–2012, 249 in 2012–2013 and 238 in 2013–2014. On cross-checking the data, we found one Trust that self-reported 12 Never Events in one year, whereas the official record states that there were two. The Trust clarified that it had 'one Never Event relating to 'wrong implant/prosthesis in ophthalmic surgery' that involved 11 patients.' We have treated this as a single Never Event.

Information on Never Event category was unavailable for 18 cases during 2011–2013. The degree of harm was reported for 429 Never Events (210 in 2011–2012 and 219 in 2012–2013; Table 1); the severe harm category included two deaths. The 'wrong implant/prosthesis' Never Events comprised 29 (43%) ophthalmic wrong lens, 10 (15%) orthopaedic (seven hip, one knee, one not categorised) and 28 (42%) not categorised. The 'retained objects' Never Events comprised 50 (18%) retained vaginal swabs, six (2%) retained throat packs, and 228 (80%) unclassified objects (including surgical instruments, guidewires, swabs, and portions of surgical glove).

**Table 1** Categories of Never Event and reported degree of harm during the period 2011–2013. Values are number (proportion).

| Never Event category | No harm | Low harm | Moderate harm | Severe harm or death | Unknown | Total |
|---|---|---|---|---|---|---|
| Wrong site surgery | 19 (14.1%) | 47 (34.8%) | 44 (32.6%) | 17 (12.6%) | 8 (5.9%) | 135 (28%) |
| Wrong implant/prosthesis | 15 (22.4%) | 22 (32.8%) | 25 (37.3%) | 3 (4.5%) | 2 (3.0%) | 67 (14%) |
| Retained foreign object | 69 (24.3%) | 66 (23.2%) | 88 (30.9%) | 14 (4.9%) | 47 (16.5%) | 284 (58%) |
| Total | 103 (21.2%) | 135 (27.8%) | 157 (32.3%) | 34 (7.0%) | 57 (11.7%) | 486 |

 

There were no differences in incidence of Never Events according to year (Fig. 1; Table 2). Table 3 gives the risks for all Never Events, surgical Never Events by category, and risk of severe harm or death according to the number of operations performed. Recalculating the overall risk after including the extra 10 patients affected by one Never Event (see above) to give a total of 752 changes the point estimate of risk only minimally to 1 in 16 204 (15 087–17 404).

There was a correlation between the number of operations and the number of operating theatres in trusts ($R^2$ 0.76; $p < 0.0001$), with approximately 1230 operations per operating theatre per year. The frequency distribution of Never Events according to number of operating theatres is shown in Fig. 2, and the risk estimates are shown in Table 4. Recalculating the overall risk after including the extra 10 patients, as above, to give a total of 752 changes the point estimate of risk to 1 in 12.8 (11.9–13.7).

There was no significant association between number of Never Events at an organisation in any one of the three years and the number of Never Events in either of the other two years when organisation size was taken into account. There was no significant association between organisation size and SHMI (number of theatres: $p = 0.07$; number of operations: $p = 0.23$). SHMI had a small negative association with the number of Never Events reported; this was consistent across all of the models and was independent of surgical caseload ($R^2$ 0.12–0.45; $p < 0.0001$; Appendix 2).

Comparative data for Never Event risks from other published sources are shown in Table 5 [18–29].

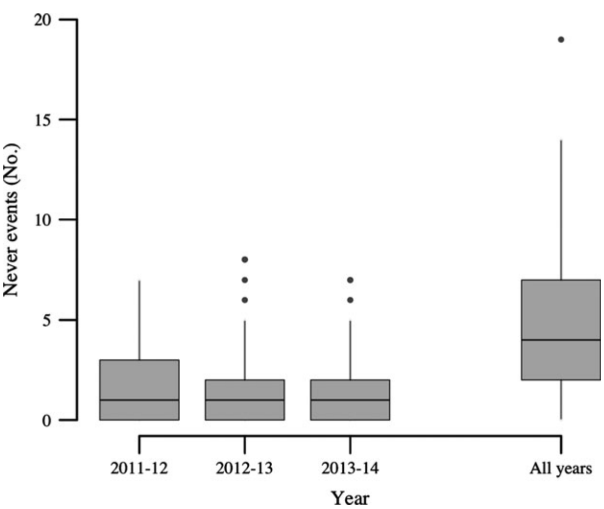


**Figure 1** Boxplot of number of surgical Never Events in each Trust according to year of reporting. Horizontal line is median, box is 25th and 75th quartiles, whiskers are 1.5 × IQR, dots are outliers.

**Table 2** Estimated risks of surgical Never Events in English NHS acute trusts according to number of operations or number of operating theatres for all three years combined and for individual years.

| Outcome | Events/ denominator | Risk 1 in N (95% CI) |
|---|---|---|
| Surgical Never Events, operations | 742/12 185 883 | 16 423 (15 283–17 648) |
| 2011–12 | 255/4 110 765 | 16 121 (14 259–18 224) |
| 2012–13 | 249/4 013 157 | 16 117 (14 235–18 247) |
| 2013–14 | 238/4 061 961 | 17 067 (15 032–19 377) |
| Surgical Never Events, operating theatres | 742/9600* | 12.9 (12.1–13.9) |
| 2011–12 | 255/3200 | 12.5 (11.2–14.1) |
| 2012–13 | 249/3200 | 12.9 (11.4–14.5) |
| 2013–14 | 238/3200 | 13.4 (11.9–14.5) |

*The number of operating theatres is 3200. This has been multiplied by three to provide an annual rate of Never Events per theatre per year

### *Modelling*

There was significant over-dispersion of the data and the AIC was lower for the negative binomial models (AIC: number of operating theatres 707; number of operations 731) than the Poisson model (AIC: number of operating theatres 712; number of operations 749) for all models tested; therefore, results are presented solely for the negative binomial model. Using number of operating theatres as a predictor resulted in a lower AIC than using number of operations reported by Trusts. Including SHMI in any of the models resulted in inconsistent, and relatively small, improvements in AIC (see Appendix 2). SHMI is only available for non-specialist Trusts. Given the marginal benefit of including it in the model and the lower generalisability, the results are therefore reported with only number of operations and number of operating theatres as predictors. Modelling

 

**Table 3** Estimated risk of surgical Never Events according to number of operations performed and category. Values are number (95% CI).

| Outcome | Never Events (n) | Risk |
|---|---|---|
| Surgical Never Events (2011–2014); 12 185 883 operations | 742 | 1 in 16 423 (15 283–17 648) |
| Surgical Never Events (2011–2013); 8 123 922 operations | 504 | 1 in 16 119 (14 771–17 588) |
| Never Events causing severe harm or death | 34 | 1 in 238 939 (170 996–333 878) |
| Wrong site surgery | 135 | 1 in 60 177 (50 846–71 220) |
| Wrong site surgery causing severe harm or death | 17 | 1 in 477 878 (298 378–765 363) |
| Wrong implant/ prosthesis | 68 | 1 in 119 749 (94 249–151 439) |
| Wrong implant/ prosthesis causing severe harm or death | 3 | 1 in 2 707 974 (920 956–7 962 516) |
| All retained foreign objects | 284 | 1 in 28 605 (25 466–32 131) |
| Retained foreign objects excluding vaginal swabs after delivery | 234 | 1 in 34 718 (30 545–39 460) |
| Retained foreign objects causing severe harm or death | 14 | 1 in 580 280 (345 676–974 108) |

using other regression techniques [15] gave almost identical results (Appendix 1).

None of the Safety Metrics were significant predictors of the risk of Never Events either alone, or combined with models including number of theatres (univariate: CQC – p = 0.41, staffing – p = 0.87, staff recommendation – p = 0.21, open and honest – p = 0.22; multivariate: CQC – p = 0.91, staffing – p = 0.30, staff recommendation – p = 0.42, open and honest – p = 0.23). However, due to the recent introduction of these metrics, data were limited (CQC – n = 55, staffing – n = 150, staff recommendation – n = 150, open and honest – n = 135).

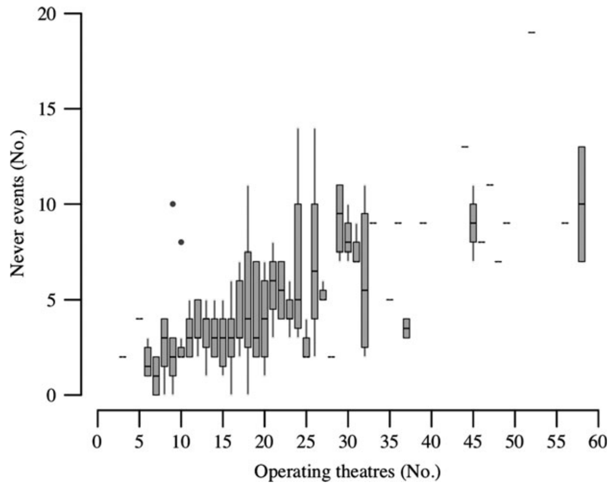


**Figure 2** Boxplot of surgical Never Events in 2011–2014 according to number of operating theatres in a trust. Horizontal line is median, box is 25th and 75th quartiles, whiskers are 1.5 × IQR, dots are outliers.

**Table 4** Estimated risk of surgical Never Events per operating theatre per year as the denominator. Values are number (95% CI).

| Outcome | Never Events (n) | Risk |
|---|---|---|
| Surgical Never Events (2011–2014); 3200 operating theatres | 742 | 1 in 12.9 (12.1–13.9) |
| Surgical Never Events (2011–2013); 3200 operating theatres | | |
| Wrong site surgery | 135 | 1 in 47 (40–56) |
| Wrong implant/ prosthesis | 68 | 1 in 96 (75–121) |
| All retained foreign objects | 284 | 1 in 23 (20–25) |
| Retained foreign objects excluding vaginal swabs after delivery | 234 | 1 in 27 (24–31) |

The estimated coefficients and residual deviance of the two models are shown in Table 6. Model fit was adequate for all three years combined and for individual years. Thus, for every additional theatre, the expected number of Never Events increases by just over 3%; or for every 10 000 operations the expected number of Never Events increases by just under 7%.

Graphical representations of the observed, predicted and 95% CI for individual trusts based on

 

**Table 5** Reported estimates of Never Events from international sources. Values are number (95% CI).

| Data source | Outcome | Methods – denominator | Population | Reported risk | Risk |
|---|---|---|---|---|---|
| Surgical Never Events | | | | | |
| NHS England [9] | All surgical Never Events | Data extraction from NRLS and STEIS; estimated number of operative cases* | All England NHS cases; 1 year | 272/3 961 497 | 1 in 14 564 (12 933–16 401) |
| Wrong site surgery | | | | | |
| NHS England [9] | All wrong site surgery | Data extraction from NRLS and STEIS; estimated number of operative cases | All England NHS cases; 1 year | 70/3 961 497 | 1 in 56 593 (44 798–71 494) |
| Robinson and Muir [18] | Wrong site surgery | NPSA and NHSLA; HES data – all admissions | All England and Wales; 2001–2006 | 218/45 369 121 | 1 in 208 115 (182 262–237 636) |
| Kwaan et al. [19] | Wrong site surgery | Insurance database; surgical procedure volume from American Hospital Association | ≈1/3 of Massachusetts hospitals; 20 years | 40/2 826 367 | 1 in 70 659 (51 894–96 210) |
| Niely et al. [20] | Wrong site surgery and wrong implants | VHA hospitals; VHA National Center for Patient Safety database. Events occurring in operating room. | All 130 VHA hospitals; 2001–2006 | 108/2 028 233 | 1 in 18 780 (15 556–22 672) |
| Niely et al. [21] | Wrong site surgery and wrong implants | VHA hospitals; VHA National Center for Patient Safety database. Events occurring in operating room | All 130 VHA hospitals; 2006–2009 | 50/1 250 000† | 1 in 25 000 (18 965–32 956) |
| Retained objects | | | | | |
| NHS England | All retained foreign objects | Data extraction from NRLS and STEIS; Estimated number of operative cases | All England NHS cases; 1 years | 161/3 961 497 | 1 in 24 606 (21 087–28 711) |
| Cima et al. [22] | All retained foreign objects | Reporting system within single institution; all patients x-rayed immediately postoperatively | All operations within single hospital 2003–2006 | 34/191 168 | 1 in 5623 (4024–7856) |
| Egorova et al. [23] | All retained foreign objects | Reporting system of institution; coronary artery bypass graft surgery | Major academic healthcare centre with affiliated hospitals | 22/153 263 | 1 in 6967 (4601–10 548) |

(continued)

 

**Table 5** (*continued*)

| Data source | Outcome | Methods – denominator | Population | Reported risk | Risk |
|---|---|---|---|---|---|
| Bani-Hani et al. [24] | All retained foreign objects | Patients re-presenting to hospitals | Irbid Province, Jordan; 1990–2003 | 11/55 300 | 1 in 5027 (2808–9003) |
| Gawande et al. [25] | All retained foreign objects | Insurance database | 22 hospitals in Massachusetts; 1985–2001 | 54/NA | NA |
| Rosen et al. [26] | All retained foreign objects | Agency for Healthcare Research and Quality Patient Safety Indicators; discharge codes | VHA Hospitals, all acute admissions; 2000–2001 | 73/430 536 | 1 in 5898 (4691–7415) |
| Zhan et al. [27] | All retained foreign objects | Agency for Healthcare Research and Quality Patient Safety Indicators; discharge codes; all admissions | 994 acute care hospitals; 2000 | 536/6 572 845 | 1 in 12 263 (11 268–13 346) |
| Romano et al. [28] | All retained foreign objects | Agency for Healthcare Research and Quality Patient Safety Indicators; discharge codes; surgical patients 'at risk' | 994 acute care hospitals; 1995–2000 | 2284/9 516 666 | 1 in 4167 (3999–4341) |
| Wrong implant | | | | | |
| Kelly and Jalil [29] | Wrong lens | Data extraction from NRLS; Number of primary procedures including operation on lens from HES | All England; 2003–2010 | 164/3 062 300 | 1 in 18 673 (16 025–21 757) |

*The number of operative cases has been estimated by multiplying the number of surgical Finished Consultant Episodes (FCE) by 0.59, which is the mean ratio between Hospital Episode Statistics recorded FCE and operative cases reported by trusts during this study.

†Denominator data for Niely et al. [21] were estimated using extrapolation and back calculation from the rates given in Niely et al. [20].

NRLS, National Reporting and Learning Service; STEIS, Strategic Executive Information System; NPSA, National Patient Safety Agency; NHSLA, NHS Litigation Authority; HES, Hospital Episode Statistics; VHA, Veterans Health Administration.

number of operations are shown in Fig. 3 and by number of theatres in Fig. 4. There were six Trusts that had more Never Events than the predicted 95% confidence limit using either number of operations or number of theatres. Of these, four fell outside these limits in a single year. Two Trusts had less Never Events than the predicted 95% confidence limit using either number of operations or number of theatres; one of these fell below these limits in a single year.

## Discussion

This analysis of Never Event data provides an estimate of the risk of surgical Never Events in England based on surgical caseload. Our data are novel in providing a per-organisation analysis and a robust statistical model of the expected number of surgical Never Events. The data support the hypothesis that Never Events should be viewed as rare, random events. Never Events are not related in a meaningful way to the SHMI index of

 

**Table 6** Results of negative binomial prediction model using number of operations and number of operating theatres as predictors for number of Never Events.

| | Intercept | Coefficient | $e^{coefficient}$ (95% CI) | p |
|---|---|---|---|---|
| Number of operations | | | | |
| All (2011–2014) | 0.889 | $7.75 \times 10^{-6}$ | 1.000008 (1.000006–1.00001) | 0.39 |
| 2011–2012 | −0.071 | $6.57 \times 10^{-6}$ | 1.000020 (1.000011–1.000029) | 0.09 |
| 2012–2013 | −0.325 | $9.04 \times 10^{-6}$ | 1.000027 (1.000017–1.000036) | 0.12 |
| 2013–2014 | −0.193 | $7.11 \times 10^{-6}$ | 1.000021 (1.000013–1.000029) | 0.14 |
| Number of operating theatres | | | | |
| All (2011–2014) | 0.801 | 0.0329 | 1.033 (1.027–1.040) | 0.37 |
| 2011–2012 | −0.183 | 0.0297 | 1.030 (1.018–1.042) | 0.09 |
| 2012–2013 | −0.411 | 0.0375 | 1.038 (1.027–1.049) | 0.09 |
| 2013–2014 | −0.302 | 0.0314 | 1.031 (1.022–1.041) | 0.10 |

p, Chi-squared test of residual deviance of model.

hospital mortality, nor to any routinely reported metric of patient safety. There is no strong evidence of more than one population of hospitals with regard to Never Event risk, although we cannot exclude this possibility. The risk of serious harm from surgical Never Events in England is very low but not zero.

In some respects, our findings are no surprise – other things being equal, larger organisations should have more Never Events. However, we suggest that this is important for three reasons. First, although, *post hoc*, this is the 'expected' answer, it has not been formally described and medicine (and politics) has repeatedly made the mistake of assuming that things are self-evidently true without checking. Second, for the Trusts and regulators, it may provide some context as to whether the number of Never Events occurring in a Trust is particularly high (or low). Third, size appears to be the only thing that matters – metrics of safety appear to be completely unrelated.

We believe our data are relatively robust, although we acknowledge the imprecision of the numerator and denominator data. A small number of English Never Events will have occurred in private hospitals and small providers not covered by our sample. These missing data are unlikely to influence the interpretation of our data.

There is no definitive record of the number of operations performed in the UK so we have chosen to

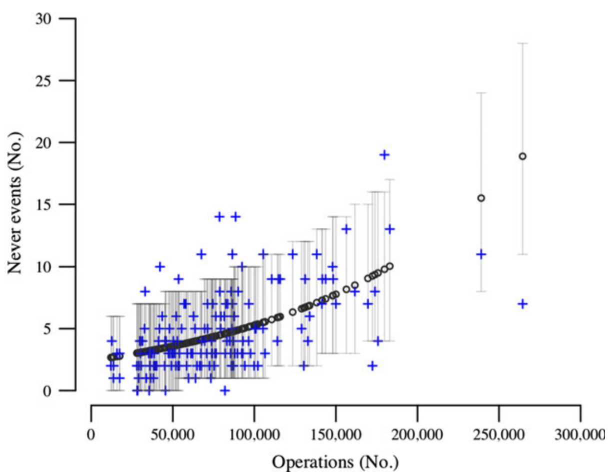


**Figure 3** Predicted and reported number of surgical Never Events according to number of operations per trust performed during 2011–2014. Error bars are predicted 95% CI.○ Predicted + Reported.

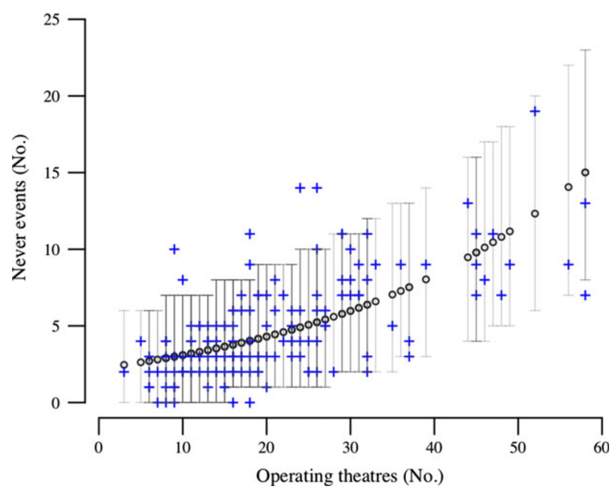


**Figure 4** Predicted and reported number of surgical Never Events according to number of operating theatres per Trust during 2011–2014. Error bars are predicted 95% CI.○ Predicted + Reported.

present the data using two estimates of surgical workload. The number of operative cases recorded by each Trust is assumed to be a relatively robust estimate of true surgical caseload and agrees with other recent estimates using different methodology. Previous estimates of UK surgical workload are similar [30]. At individual Trust level, the results are qualitatively the same using both number of recorded operative cases and number of operating theatres as estimates of surgical caseload; there is a high correlation between the two, with an average number of cases per theatre per year of around 1230. Given the ease with which organisations can simply count the number of operating theatres, compared with the difficulty some Trusts (and the NHS) had in returning a reliable number of operations performed, we would suggest that the number of operating theatres is a better measure of surgical workload. It is easily measured and possibly has more meaning to clinicians and regulators.

There is no single metric of patient safety that we could use confidently. Standardised mortality ratios are controversial [31, 32], and are certainly too blunt an instrument to describe safety on their own. Vincent et al. [10] argue for a suite of complementary safety metrics addressing past harm, reliability of behaviour and systems, sensitivity to operations (near real-time safety monitoring), anticipation and preparedness and integration-and-learning (responding to and learning from safety information). If such metrics had been available, then we would have used them. The NHS Choices metrics are a limited subset of the above, providing crude information on staffing (reliability of systems), staff perceptions of their own organisation (global rating of quality) and a composite perception of organisational incident reporting and management (integration-and-learning). The CQC inspection safety rating is another global safety assessment based on a whole suite of safety metrics as well as direct inspection of hospitals. None of these metrics showed any association with the number of Never Events. Given the smaller number of organisations with these metrics, we cannot exclude a lack of statistical power for such an association.

The SHMI is an (imperfect) indicator of past harm and has a negative association with the number of Never Events reported. Taken at face value, the negative association between SHMI and Never Events supports our contention that Never Events have little to do with organisational safety, but we do not wish to read too much into this unexpected finding.

The actual number of Never Events is likely to be higher than reported here through failure to recognise, failure to report within the organisation, and variations in interpretation of Never Events by Trusts and regulators. Wrong site surgery is likely to be recognised at the time, but retained objects may have a significant time lag [25]. There is certainly support for the belief that the decision to label a specific event as a Never Event may not be consistent. Eleven incidents that we believe could have been classed as Never Events were declared only as serious incidents by respondent Trusts. These included wrong site surgery 'noted at the time', wrong implants 'decided not to change' and retained objects 'not sutured in'. The insertion of a wrong lens during ophthalmic surgery appears to be a particularly flexible definition; a quote from one Trust's published 'Quality Account' after such an occurrence was '. . . the ophthalmology incidents could be interpreted as outside the criterion of the Never Event, as further surgery was not required and the patients were happy with the outcome of their surgery. . .'. Currently, there are a variety of reporting mechanisms for NHS organisations to report Never Events, which NHS England is attempting to address, but failure to report is considered a major failing by regulators. There has been a significant increase in estimated event risks between 2008 [33] and 2011–2012. We would suggest that this reflects an increase in recognition and reporting of Never Events.

The risk of any particular type of Never Event is dependent on surgical practice, for instance, only hospitals providing cataract services can report a retained lens. We did not seek to subcategorise surgical practice within organisations, so we cannot exclude the possibility that certain hospitals are more at risk because of their workload pattern. However, almost all UK Trusts are single general hospitals or a mix of specialist and general hospitals. Recent data from NHS England [34] show 18 'wrong lens' Never Events in 2014–2015 out of around 375 000 cataract operations – the 1 in ~21 000 risk of a wrong lens event is therefore broadly similar to that of all surgical Never Events.

 

We have taken a straightforward approach to analysis of the data. Importantly, we have demonstrated that the distribution of Never Events is consistent with the hypothesis that they are independent random events. We found no evidence that the distribution of Never Events is different to what would be expected to occur by chance. This, to an extent, could be viewed as a contradiction to the regulatory and political view: ‘Repeated Never Events, particularly if they are the same type of incident, could demonstrate a failure of the organisation’s leadership, particularly clinical leadership, to take patient safety seriously’ [35]. However, we believe that these aspects are consistent. Never Events are random, rare events; their occurrence does not imply anything particularly unusual about an organisation. An analogy would be finding multiple national lottery winners from a single street. This is an inevitable event that occurs solely by chance and provides no information about the street occupants other than that they buy lottery tickets. Never Events, however, provide a focus to review safety culture, policies and practices. Conversely, and importantly, the lack of Never Events provides no assurance that safety culture and practice is good within an organisation. For smaller organisations, it is likely to be related to chance.

The negative binomial model provides a reasonable fit for the data. As expected from a statistical model, a small number of Trusts fell outside the 95% CI for the number of Never Events in a single year, or over the three studied years combined. Of note, of the six Trusts that fell above the 95% CI over the three years, two did not fall outside these limits in any single year. The other four were outside the limits in only one of the three years. Our data provide a model to estimate the expected number of Never Events. This may be useful for regulators and those with healthcare oversight to identify those organisations, where Never Event risks are significantly above or below expected.

The interpretation of the Never Event risk is context dependent. On the one hand, one Never Event in approximately 16 000 operations is clearly an unusual event; moderate or severe harm is even rarer. There is therefore little chance that an individual member of an operating team will ever be personally involved in a Never Event, regardless of their own individual practice. On the other hand, Never Events are not that uncommon at the level of the organisation; the probability of a median sized trust (~ 24 000 operations or 17 theatres) having no surgical Never Events over a three-year period is around 2%.

At a national as well as an individual patient level, just one Never Event is too many. Of all the Never Events, retained objects, wrong site surgery and wrong implant seem the most straightforward to prevent, relying on simple checking procedures. Our data provide some context to this risk. Patients are warned of similar frequency complications as part of the consent process: ophthalmologists warn of the risk of infective endophthalmitis following vitreoretinal surgery with an estimated risk of 1 in 10 000 [36]; anaesthetists warn of the risk of permanent complications of central neuraxial blocks with an estimated risk of 1 in 24 000 – 1 in 50 000 [37]. Recent data suggest a similar risk of unintended awareness following general anaesthesia of 1 in 19 000 [38]. The risk of serious airway complications associated with general anaesthesia is estimated to be around 1 in 22 000 general anaesthetics [39], which is approximately three times more common than wrong site surgery. The estimated risk of death or brain damage from airway complications is 1 in approximately 150 000 [39], which is again approximately three times more common than severe harm due to wrong site surgery. Accurate data on the mislabelling of surgical specimens in the UK are not available, but two US-based studies estimate the risks of wrong patient identification at around 1 in 2500 cases [40, 41].

Our data are in line with the NPSA/NHS England reported Never Events [9]. The slightly lower risk in our data probably reflects the small number of surgical Never Events that occur outside the operating theatre; these would have been captured by the NPSA and NHS England but not in our sample. The risk of wrong site/wrong implant Never Events is similar to that reported by Kwaan et al. [19] and Neily et al. [20, 21] from the USA. All trusts in the UK declared that they have implemented the World Health Organization (WHO) Checklist before these incidents occurred. The incidents reported by Kwaan et al. [19] were collected predominantly before the introduction of the ‘Universal Protocol’ in the USA, which contains an almost identical checking process before surgery as the WHO checklist [42]. The more prolonged study of incorrect surgical procedures from the Veterans Health Administration hospitals in the USA [21] concluded that the reduction in event risk was associated with the introduction of medical team training, which has a strong emphasis on pre- and postoperative briefings, checklists and teamwork, similar to the UK 'Five steps to safer surgery' [43]. Our data provide no information on the effectiveness of the checks within the WHO checklist, as there are no comparable data from before its introduction. However, the occurrence of these events of course means that the WHO checklist process has failed in some way for at least some of the cases. The data from Robinson and Muir [18] used NPSA (self-reported) and NHS litigation authority (legal claims) sources. The reporting culture in the UK at that time was less robust than now, and claims analysis is inevitably incomplete. In addition, the denominator for their data was all hospital admissions, rather than our estimate of operations.

Never Events are held to be sentinel events for quality of surgical care [2]. We do not expect Never Events to have an effect on mortality as they are too rare, and the risk of serious harm is too low. However, evidence from the SURPASS checklist suggests that there is a strong link between compliance with a checklist process and surgical outcome [44]. Similarly, the Veterans Health Administration Medical Team Training Program found an association between surgical outcomes and both training in the program and the degree of engagement with the process [45]. For a surgical Never Event to occur requires a variety of relatively unusual errors to occur at the same time, only one of which is a failure of the checklist process. Poor practices in the operating theatre will therefore only occasionally result in Never Events. Poor communication, planning, teamwork and a failure to check the basics are more likely result in non-Never Event related harm and surgical mortality [44]. We would suggest that the focus for front-line staff, managers and regulators interested in improving surgical safety and the use of the WHO checklist should be on common surgical outcomes and complications. We cannot exclude the possibility that there are significant variations in quality and safety of surgical care within NHS hospitals. Never Event occurrence may be related to a yet unmeasured set of safety behaviours, but Occam's razor would argue for a simpler explanation. Never Events are important, but as they are rare, apparently random events they are the wrong metric to gauge safety within the operating theatre [46].

## Acknowledgements

We thank Nazurah Wahid for assistance with cross-checking data, both internally and with external sources.

## Competing interests

No external funding and no competing interests declared. IKM has discussed the role of the WHO checklist, and the concepts in this paper, with members of the NPSA/NHS Commissioning Board/NHS England prior to publication, but those persons have had no role in the design, analysis or conclusions of this paper. Both IKM and SHM act as specialist advisors to the NHS Care Quality Commission.

## References

1. National Patent Safety Agency. Never Events Annual Report 2009/10, 2010. http://www.nrls.npsa.nhs.uk/resources/?entryid45 = 83319 (accessed 16/09/2015).
2. Michaels RK, Makary MA, Dahab Y, et al. Achieving the National Quality Forum's ''Never Events''. Prevention of wrong site, wrong procedure, and wrong patient operations. *Annals of Surgery* 2007; **245**: 526–32.
3. Department of Health. The ''never events'' list for 2011/12, 2011. https://www.gov.uk/government/publications/the-never-events-list-for-2011-12 (accessed 16/09/2015).
4. Monitor. NHS foundation trusts found in significant breach of their authorisation (Cambridge University Hospitals NHS Foundation Trust), 2012. http://webarchive.nationalarchives.gov.uk/20140106173514/http://monitor-nhsft.gov.uk/about-your-local-nhs-foundation-trust/regulatory-action/nhs-foundation-trusts-found-significant-brea (accessed 16/09/2015).
5. The Guardian. NHS regulator intervenes after botched operations at Cambridge hospitals, 2012. http://www.theguardian.com/society/2012/nov/23/nhs-botched-operations-cambridge-hospitals (accessed 16/09/2015).
6. Mail Online. Maimed by the NHS: Catalogue of errors revealed as patients are KILLED by drugs meant for others, 2012. http://www.dailymail.co.uk/health/article-2122872/Maimed-NHS-Catalogue-errors-revealed-patients-KILLED-drugs-meant-others.html (accessed 16/09/2015).
7. NHS England Patient Safety Domain. Standardise, educate, harmonise. Commissioning the conditions for safer surgery. Report of the NHS England Never Events Taskforce, 2014. http://www.england.nhs.uk/wp-content/uploads/2014/02/sur-nev-ev-tf-rep.pdf (accessed 20/05/2015).
8. NHS England. Seven steps to patient safety, 2014. www.npsa.nhs.uk/sevensteps/(accessed 20/05/2015).

 

9. NHS England. Never Events Data, 2015. http://www.england.nhs.uk/ourwork/patientsafety/never-events/ne-data (accessed 17/09/2015).
10. Centre for Patient Safety & Service Quality. The measurement and monitoring of safety, 2013. http://www.health.org.uk/publication/measurement-and-monitoring-safety (accessed 17/09/2015).
11. Health and Social Care Information Centre. Indicator Portal, 2015. https://indicators.ic.nhs.uk/webview/(accessed 17/09/2015).
12. University of Sheffield. An evaluation of the Summary Hospital Mortality Index, 2011. http://www.sheffield.ac.uk/polopoly_fs/1.51777!/file/SHMI_Final_Report.pdf (accessed 17/09/2015).
13. NHS Choices. Results for all patients nationwide, 2015. http://www.nhs.uk/Service-Search/All-patients/Nationwide/Results/Accountability/646%2c998/0/0/1856/0?orgTypes=GPB%2CHOS&distance=25&nationwide=True&tabText=Hospital&metricGroupId=0&metricGroupText=All%20patients&resultsOnPageValue=100 (accessed 17/09/2015).
14. The R Foundation. The R Project for Statistical Computing, 2015. http://www.R-project.org/(accessed 17/09/2015).
15. Wirtschafts Universitat Wein. Regression Models for Count Data in R, 2007. http://epub.wu.ac.at/1168/(accessed 17/09/2015).
16. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 1974; **19**: 716–23.
17. Wilson EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 1927; **22**: 209–12.
18. Robinson PM, Muir LT. Wrong-site surgery in orthopaedics. *Bone and Joint Journal* 2009; **91-B**: 1274–80.
19. Kwaan MR, Studdert DM, Zinner MJ, Gawande AA. Incidence, patterns, and prevention of wrong-site surgery. *Journal of the American Medical Association Surgery* 2006; **141**: 353–8.
20. Neily J, Mills PD, Eldridge N, et al. Incorrect surgical procedures within and outside of the operating room. *Journal of the American Medical Association Surgery* 2009; **144**: 1028–34.
21. Neily J, Mills PD, Eldridge N, et al. Incorrect surgical procedures within and outside of the operating Room: a follow-up report. *Archives of Surgery* 2011; **146**: 1235–9.
22. Cima RR, Kollengode A, Garnatz J, Storsveen A, Weisbrod C, Deschamps C. Incidence and characteristics of potential and actual retained foreign object events in surgical patients. *Journal of the American College of Surgeons* 2008; **207**: 80–7.
23. Egorova NN, Moskowitz A, Gelijns A, et al. Managing the prevention of retained surgical instruments: what is the value of counting? *Annals of Surgery* 2008; **247**: 13–8.
24. Bani-Hani KE, Gharaibeh KA, Yagha RJ. Retained surgical sponges (gossypiboma). *Asian Journal of Surgery* 2005; **28**: 109–15.
25. Gawande AA, Studdert DM, Orav EJ, Brennan TA, Zinner MJ. Risk factors for retained instruments and sponges after surgery. *New England Journal of Medicine* 2003; **348**: 229–35.
26. Rosen AK, Rivard P, Zhao S, et al. Evaluating the patient safety indicators: how well do they perform on Veterans Health Administration data? *Medical Care* 2005; **43**: 873–84.
27. Zhan C, Miller MR. Excess length of stay, charges and mortality attributable to medical injuries during hospitalization. *Journal of the American Medical Association* 2003; **290**: 1868–74.
28. Romano PS, Geppert JJ, Davies S, Miller MR, Elixhauser A, McDonald KM. A national profile of patient safety in U.S. hospitals. *Health Affairs* 2003; **22**: 154–66.
29. Kelly SP, Jalil A. Wrong intraocular lens implant; learning from reported patient safety incidents. *Eye* 2011; **25**: 730–4.
30. Woodall NM, Cook TM. National census of airway management techniques used for anaesthesia in the UK: first phase of the Fourth National Audit Project at the Royal College of Anaesthetists. *British Journal of Anaesthesia* 2011; **106**: 266–71.
31. Mohammed MA, Deeks JJ, Girling A, et al. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *British Medical Journal* 2009; **338**: b780.
32. Lilford R, Pronovost P. Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. *British Medical Journal* 2010; **340**: c2016.
33. Panesar SS, Noble DJ, Mirza SB, et al. Can the surgical checklist reduce the risk of wrong site surgery in orthopaedics? - can the checklist help? Supporting evidence from analysis of a national patient incident reporting system. *Journal of Orthopaedic Surgery and Research* 2011; **6**: 18.
34. NHS England. Provisional publication of Never Events reported as occurring between 1 April 2014 and 31 March 2015, 2015. http://www.england.nhs.uk/wp-content/uploads/2015/07/provsnl-ne-data-04-14-to-03-15.pdf (accessed 17/09/2015).
35. NHS England. The never events policy framework: an update to the never events policy, 2012. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213046/never-events-policy-framework-update-to-policy.pdf (accessed 17/09/2015).
36. Chu DS, Foster CS. Sympathetic ophthalmia. *International Ophthalmology Clinics* 2002; **42**: 179–85.
37. Cook TM, Counsell D, Wildsmith JAW. Major complications of central neuraxial block: report on the Third National Audit Project of the Royal College of Anaesthetists. *British Journal of Anaesthesia* 2009; **102**: 179–90.
38. Pandit JJ, Andrade J, Bogod DG, et al. The 5th National Audit Project (NAP5) on accidental awareness during general anaesthesia: summary of main findings and risk factors. *Anaesthesia* 2014; **69**: 1089–101.
39. Cook TM, Woodall N, Frerk C. Major complications of airway management in the UK: results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: Anaesthesia. *British Journal of Anaesthesia* 2011; **106**: 617–31.
40. Makary MA, Epstein J, Pronovost PJ, Millman EA, Hartmann EC, Freischlag JA. Surgical specimen identification errors: a new measure of quality in surgical care. *Surgery* 2007; **141**: 450–5.
41. Nakhleh RE, Idowu MO, Souers RJ, Meier FA, Bekeris LG. Mislabeling of cases, specimens, blocks, and slides. A College of American Pathologists study of 136 institutions. *Archives of Pathology and Laboratory Medicine* 2011; **135**: 969–74.
42. Haynes AB, Weiser TG, Berry WR, et al. A Surgical Safety Checklist to reduce morbidity and mortality in a global population. *New England Journal of Medicine* 2009; **360**: 491–9.
43. NHS England. 10 for 2010: Five steps to safer surgery, 2010. http://www.nrls.npsa.nhs.uk/resources/collections/10-for-2010/five-steps-to-safer-surgery/ (accessed 17/09/2015).
44. de Vries EN, Prins HA, Crolla RMPH, et al. Effect of a comprehensive surgical safety system on patient outcomes. *New England Journal of Medicine* 2010; **363**: 1928–37.
45. Neily J, Mills PD, Young-Xu Y, et al. Association between implementation of a medical team training program and surgical mortality. *Journal of the American Medical Association* 2010; **304**: 1693–700.
46. Harrop-Griffiths W. Never Events. *Anaesthesia* 2011; **66**: 158–62.

## Appendix Alternative modelling approaches

The Poisson model assumes that the variance of the sample approximates to the mean. This is rarely the case in biological data, and commonly the variance is greater than the mean (over-dispersion). Two related effects are seen; first, the confidence in the predicted values is over-stated, and second, the Poisson model tends to under-predict the number of zero counts. Various methods can be used to address this. Confidence intervals can be adjusted to more robust estimates using the sandwich covariance matrix estimator or the dispersion parameter can be unrestricted (i.e. variance $\not\approx$ mean) (quasiPoisson). Alternatively, different distributions can be used such as the negative binomial. Third, composite distributions can be used which provide separate estimates of a binary outcome (e.g. Never Event or not) and the count of such events. Examples of these include zero-inflated negative binomial models and hurdle models. A fuller theoretical and empirical analysis is given by Zeileis et al. [15].

To undertake a sensitivity analysis of the data, we re-analysed the main dataset (all surgical Never Events predicted by number of operations or number of operating theatres). The data are presented below (Tables A1 and A2). Analysis of the yearly event data demonstrated qualitatively similar results (data not shown). The Poisson model performed least well, with a greater AIC and under-prediction of the zero counts (This effect is more marked in the single year data where zero counts are higher). The adjusted (sandwich) and quasiPoisson models give the same mean estimates but, as expected wider standard errors. There is little to choose between the negative binomial, zero-inflated negative binomial and hurdle models. Without any empirical evidence to suggest that there is a different process determining whether a Never Event occurs from how many are there, we feel that Occam's razor should be invoked and the simpler, negative binomial model used.

## Appendix Never Events and Summary Hospital Mortality Index (SHMI) data.

SHMI has a negative association with the number of Never Events reported both with simple linear regression and with negative binomial regression (Table A3). This is consistent across all of the models and is independent of surgical caseload. In other words, as SHMI increases ('worse' outcomes) the number of reported Never Events declines. Trusts with a lower (better) SHMI report more Never Events. We are, however, at somewhat of a loss to explain this finding. Broadly, there are three non-exclusive explanations: 1. The relationship is real: organisations with fewer Never Events have more than their expected number of deaths. This seems implausible. 2. There is an issue with the risk adjustment model underlying SHMI. SHMI covers all deaths and it is not possible to disentangle surgical deaths. Of note, there is a weak (non-significant) association between SHMI and size – SHMI decreases as size increases. This holds true for the data published as part of the NHS transparent reporting of SHMI (all included deaths) and our data (number of operations, number of operating theatres). In other words, bigger hospitals have a slightly lower SHMI on average, so have a 'double whammy' of larger size and lower

**Table A1** Coefficients (SE) and measures of model performance for alternate modelling approaches for the overall data set with the number of operations as predictor.

| | Poisson | Adjusted Poisson (sandwich) | Quasipoisson | Negative binomial | Hurdle | Zero-inflated negative binomial |
|---|---|---|---|---|---|---|
| Intercept | 0.943 (0.074) | 0.943 (0.10) | 0.943 (0.093) | 0.887 (0.095) | 0.878 (0.048) | * |
| Number of operations ($\times 10^{-6}$) | 7.1 (0.68) | 7.1 (1.09) | 7.1 (0.86) | 7.7 (0.95) | 7.8 (NE) | * |
| AIC | 748.4 | – | – | 730.1 | 733.4 | * |
| Observed zero count | 5 | 5 | 5 | 5 | 5 | * |
| Predicted zero count | 3 | – | – | 5 | 5 | * |

*The zero-inflated negative binomial model fails to converge for number of operations.
NE, not estimable.

**Table A2** Coefficients (SE) and measures of model performance for alternate modelling approaches for the overall data set with number of operating theatres as predictor.

| | Poisson | Adjusted Poisson (sandwich) | Quasipoisson | Negative binomial | Hurdle | Zero-inflated negative binomial |
|---|---|---|---|---|---|---|
| Intercept | 0.829 (0.0758) | 0.829 (0.085) | 0.8291 (0.0878) | 0.800 (0.0877) | 0.8025 (0.0971) | 0.8078 (0.0943) |
| Number of operating theatres | 0.0317 (0.0026) | 0.0317 (0.003) | 0.0318 (0.0030) | 0.0329 (0.0032) | 0.0329 (0.0035) | 0.0328 (0.0034) |
| AIC | 710.9 | – | – | 705.9 | 709.7 | 709.9 |
| Observed zero count | 5 | 5 | 5 | 5 | 5 | 5 |
| Predicted zero count | 3 | – | – | 5 | 5 | 5 |

**Table A3** Negative binomial modelling of Never Events using surgical caseload ±SHMI (Summary Hospital Mortality Index) as predictors. The models are for all years (2011–2014). Qualitatively similar results are found for individual years though the improvement in AIC is inconsistent. Note the data set used is smaller than for the main results modelling as SHMI is not available for every trust (n = 140); modelled coefficients are therefore slightly different. Values are model coefficients (SE).

| | Operations | Operations + SHMI | Operating theatres | Operating theatres + SHMI |
|---|---|---|---|---|
| Intercept | 1.01 (0.11) | 3.00 (0.51) | 0.90 (0.099) | 2.62 (0.48) |
| Number of operations ($\times 10^{-6}$) | 6.37 (1.05) | 5.78 (0.98) | – | – |
| Number of theatres | – | – | 0.028 (0.003) | 0.026 (0.0034) |
| SHMI | – | −1.96 (0.49) | – | −1.67 (0.46) |
| AIC | 660.5 | 647.9 | 641.4 | 631.4 |

SHMI. 3. There is a common factor between high SHMI and low reported Never Events of under-reporting. This is plausible, although this is pure speculation on our part.

We have not included SHMI in our main reported model results for the following reasons: 1. SHMI is only applicable to non-specialist Trusts, so the result is less generalisable. 2. The improvement in AIC is marginal, and not present at all in some of the models for the individual years. Hence, we believe that it adds little value to the models presented.

 