PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Comparison of two methods to estimate adverse events in the IBEAS STUDY (Iberoamerican study of adverse events): cross-sectional versus retrospective cohort design.
AUTHORS	LOPEZ FRESNEÑA, NIEVES; ARANAZ ANDRÉS, JESÚS MARIA; LIMÓN RAMÍREZ, RAMÓN; AIBAR REMÓN, CARLOS; GEA VELAZQUEZ DE CASTRO, MARIA TERESA; Bolumar, Francisco; Hernandez-Aguado, Ildefonso; DIAZ AGERO PEREZ, CRISTINA; TEROL GARCÍA, ENRIQUE; Michel, Philippe; Sousa, Paulo; LARIZGOITIA JAUREGUI, ITZIAR;

VERSION 1 - REVIEW

REVIEWER	Natasha Rafter Royal College of Surgeons in Ireland Ireland
REVIEW RETURNED	18-Sep-2016

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. The paper provides an important comparison between two methods of adverse event estimation. The authors should define and justify the use of the prevalence study and incidence study nomenclature. Prevalence and incidence are not defined in the manuscript or referenced. The authors should consider re-naming their study designs in line with Michel et al BMJ 2004 'Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals'. This is a critical reference in this area of research but is absent from the reference list; it should be referred to in the introduction and/or discussion. Michel et al compared adverse event rates discovered via three methods – cross-sectional, prospective and retrospective. In addition, other retrospective national AE studies have calculated both incidences and prevalences so these terms do not necessarily indicate a study method. In the results text the authors refer to a one-day study and a retrospective review, these may be better terms for the text and tables than prevalence and incidence. More detail is required about the IBEAS one-day and retrospective record review studies so that the paper reads well in itself and the reader does not need to access the IBEAS references for basic details. Full descriptions of the methods employed in both studies are required in order to assess the comparison. Were the reviewers the same? If not did they receive the same training? Were the data collection forms the same? How was preventability assessed in each study? How was an adverse event defined and was it the same for

	both studies? How did the one-day study assess adverse event impact on readmission? In the retrospective study the reviewers looked for AEs 'during the entire hospitalization' – did this include AEs that occurred prior to hospitalization also (these were included in the one-day study)? The paper would also benefit from proof reading the written English. Statistics – this section describes prevalence and incidence being estimated through logistic regression taking into account covariates. However, the abstract presents crude rates. It should be clearer in the abstract and results which figures are crude and which are adjusted. Confidence intervals should be presented for all estimates in the text and tables. Results - the figures presented in the abstract should also appear in the text of the results section. Discussion - the adverse event rate of 28.9% is significantly above all other national/large adverse events studies. However the methods description is incomplete and it is difficult to be certain that the methods used are comparable to other studies. If they are then this is a very significant result and should be further discussed. The current discussion is unclear and lacks any referencing. Information on resources used in both types of study would be interesting (cost, numbers of reviewers, time taken). Specific comments: 1. Abstract – results section: explain what the range refers to. Confidence intervals should be provided. 2. Summary box – the first sentence seems incomplete 'The identification of adverse events is the first step to improve.' 3. Introduction – paragraph two has only one citation for the first sentence and its subsequent statements are inadequately referenced. 4. Introduction – paragraph three sentence one should be 'consists of' (not in). Please provide a reference in English for this important statement introducing cross-sectional surveys. 5. The definition of an adverse event used in IBEAS should be provided and stated whether this was the same for both studies. 6. Methods - pg 5, line 25. Provide an exact sample size. 7. Pg 5, line 27 'An AE was considered prevalent..' should this be present? 8. Results – the results of the full IBEAS study should be presented in the text as well as those of the one-day study. 9. Pg 7 line 6 – there were 215 patients where the incident was caused by healthcare management but there were 314 adverse events. Does this mean that more than one adverse event was able to be counted per patient? This should be made clearer in the methods and the results. 10. Pg 8 line 3 'Surgical patients were associated with more risks than patients admitted in the medical wards' – but the confidence interval for the odds ratio of the one-day study included 1 (OR 1.17 (95% CI 0.99-1.38))? 11. Pg 8 line 8 'the stay before survey' – what is this? Please define. How long prior to the admission under examination did you look? Also in table 3. 12. Pg 8 line 47 – what is 'paradigmatic'?

	13. Pg 8 last para should refer to Table 4. Because no confidence intervals are given how can one justify the conclusion that retrospective review identifies less serious and shorter adverse events etc? 14. Pg 9 line 39 – what does this mean 'Health related infections ….are considered less preventable than other types of AE and these categories are identified further in the prevalence study.'? Please explain. This information may be better suited to the discussion section. 15. Pg 9 line 42 – provide the data or state data not shown and a confidence interval for the statement 'The severity of the AE was not associated with its preventability'. 16. Discussion – pg 10 the first three paragraphs on study types require referencing. 17. Pg 10, line 25 Please explain or reference the statement 'The relationship between prevalence and incidence generally depends on the duration of the event under review and the period of observation.' This would not be the standard definition for these terms nor how they have been used in previous adverse event studies. 18. Figures 1a and 1b – label arrows. The figure notes contain definitions which should also be included in the methods. 19. Pg 11 line 31 – what do you mean by 'we will be able to detect the number of false positive through the revision of the second questionnaire (MRF2)' and 'This evaluation will enable us to adjust the alarm conditions for transversal studies'? Please explain these statements further. 20. Pg 11 line 41 - provide the references for the other AE studies referred. 21. Pg 11 last para – this should be in the methods. Nosological means classification of disease, is this the term you wish to use? 22. Pg 12 first sentence 'the reliability of the questionnaire in other studies has been assessed as moderate' – please reference. Tables – Table 1 - why was a p level of 0.1 chosen to indicate the level of nonsignificance in table 1? – Confidence intervals should be provided. – The columns are not fully labeled. – The studies should be presented consistently (eg one the left and the other on the right - table 1 differs from the others. – Table 5 – are these crude or adjusted figures? What is SD and ST?

REVIEWER	Dr Sharon Mayor Division of Population Medicine Cardiff School of Medicine Cardiff Wales UK
REVIEW RETURNED	17-Oct-2016

GENERAL COMMENTS	This is a well-conceived study addressing an important question regarding the optimal methodological approach in undertaken case record review to improve the quality of care provided. Interesting data is generated, however, I think it would be useful to see greater clarity in this paper on a number of methodological points and a little

	more clinical application and insight in the discussion. Methods Firstly, in describing the methods, it would be helpful to have a brief overview on the training provided to the reviewers. Secondly, previous studies report a modest level of inter rater reliability using the structured review forms used and if this was not assessed in this study, it should be raised as a limitation as it might explain some of the inter-country variation seen in the reported rates of AEs. Thirdly, it would be helpful to state up front in the methods that a Likert scale was used to determine the presence and preventability of detected AEs and that the threshold of 4/6 was used in both case? Fourthly, the authors state the second phase of the study took place most frequently after the patient's discharge. Can the authors clarify this proportion as the incidence study seeks to examine the whole of the inpatient episode and if a significant proportion of reviews were undertaken whilst the patient was still an inpatient the rates might be underestimated? Results On page 7 the authors state the proportion of patients experiencing an adverse event was close to 19.8%. However, they then report a total of 314 adverse events and a retrospective incidence of 28.9%. Is this as a result of multiple adverse events in individual cases or has a lower threshold been used in determining causality? Can this be clarified? The interpretation of the different rates and types of events in both studies is interesting but does highlight a number of important issues if the findings are going to be used to prioritise improvement efforts. For example, nosocomial pneumonia and neonatal complications were found to be serious AEs in the prevalence study with nosocomial pneumonia probably being of lower preventability and neonatal complications being infrequent. However, the less serious issues identified through reviewing the whole episode of care e.g UTIs and wound infections, may be more useful in identifying where improvement priorities should be focused. Discussion Can the authors be sure that the reviewers were trained enough to perform the review in a reliable way. Some countries reported rates in line with previous studies, others were three times higher than what would be expected in other settings. Serious adverse events are highlighted as being important but these are relatively infrequent. Other AEs such as skin and pressure damage and patient falls etc are unlikely to be detected in one-day prevalence surveys but offer important insights into the general quality of care provided? Should the discussion include more analysis of the relative strengths and limitations of the two designs when used for different purposes.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 Reviewer Name: Natasha Rafter Institution and Country: Royal College of Surgeons in Ireland, Ireland Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below Thank you for the opportunity to review this manuscript. The paper provides an important comparison between two methods of adverse event estimation.

1. The authors should define and justify the use of the prevalence study and incidence study nomenclature. Prevalence and incidence are not defined in the manuscript or referenced.

The cross sectional method publisher in BMJQSHC is used to estimate prevalence. The incidence is referred a refer to new cases in one-day study

2. Related to 'Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals'. This is a critical reference in this area of research but is absent from the reference list; it should be referred to in the introduction and/or discussion. Michel et al compared adverse event rates discovered via three methods – cross-sectional, prospective and retrospective.

We have added this reference:

In the same context Michel et al.[19] evaluated the rates comparing tree different methods: cross sectional, prospective and retrospective.

Michel P, Quenon JL, De Sarasqueta, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. BMJ 2004; 328(4): 199-203

3. In addition, other retrospective national AE studies have calculated both incidences and prevalences so these terms do not necessarily indicate a study method. In the results text the authors refer to a one-day study and a retrospective review, these may be better terms for the text and tables than prevalence and incidence.

We also agree with this comment, we have changed the terms in the text.

4. More detail is required about the IBEAS one-day and retrospective record review studies so that the paper reads well in itself and the reader does not need to access the IBEAS references for basic details.
Full descriptions of the methods employed in both studies are required in order to assess the comparison. Were the reviewers the same? If not did they receive the same training?
Were the data collection forms the same? How was preventability assessed in each study? How was an adverse event defined and was it the same for both studies?

We explain the methodology in the text:

The cross-sectional or prevalence study involved determining how many patients admitted to the participating hospitals experienced harmful incidents attributable to health care on a given day (day 0). A prevalent AE is defined as one that originates during hospitalization and is clinically present on the day of the study, either as an after-effect or under treatment. This also includes those AE that were occasioned prior to hospitalization at any care level and which led to subsequent admission. AE that had occurred prior to the survey and whose effects had disappeared without prolonging the hospitalization on that particular day were not included.

The retrospective incidence study was conducted using a sample of patients with the aim of confirming whether the prevalence study could replace the conventional incidence study used to date. Specifically, the study involved reviewing the case notes of a random sample of 10% of patients (1.101 patients) hospitalized on day 0, proportional to hospital size, from the entire IBEAS study population. Case notes were scanned to ascertain whether, at some point during their hospitalization (or in a previous admission to the hospital), inpatients had experienced a harmful incident, regardless of whether the consequences of the incident were still present on day 0. Patients continued to be monitored until discharge. The sampling strategy and forms are available upon request. An incident AE is defined as one that occurs during any patient care process, as it may be detected at another level of care or in other hospital. In practical terms, as we carried out a retrospective study based on clinical hospital records, primary care AE were not included. Those that led to readmission in the same or another hospital were compensated by the AE which were detected during this hospitalization and which had been originated in a previous hospitalization.

The reviewers training took place in two stages. First, the trainer workshop addressed the national coordinating teams in Buenos Aires 2007. Second, the national coordinators trained in turn the national investigators. A concordance study was carried out in Bogotá in 2008 using clinical records from each country. The most complex cases were assessed and an agreement was reached.

The preventable AE and the gravity were assessed according to the recommendations in the Modular Questionnaire, and the reviewers were also trained in these criteria.

The cross sectional and the retrospective study were made by the same reviewers in each country.

5. How did the one-day study assess adverse event impact on readmission? In the retrospective study the reviewers looked for AEs 'during the entire hospitalization' – did this include AEs that occurred prior to hospitalization also (these were included in the one-day study)? The paper would also benefit from proof reading the written English.

In both studies (cross sectional and retrospective), researchers used two tools to detect harmful incidents, namely a Screening Guide and a Modular Questionnaire[i,ii] to identify harmful incidents using the medical record review methodology[iii,iv].

First, the screening guide was applied to the patients in the study. This served as an alert and tracking system for possible incidents. All the patients admitted at hospital (except emergency room) were studied. The screening was made by well trained nurses.

If a patient screened positive for one or more of the 19 alert criteria in the screening guide, the case was studied using the case history. An in-depth study of case histories enabled researchers to conclude whether a patient did in fact present with the consequences of a harmful incident (true positive) and if so, to classify the type of event, its severity, any associated factors, and whether or not the incident could have been avoided, etc. This second confirmatory review was made (in both cross sectional and retrospective) by medical doctors with at least 5 years of clinical experience. A patient could have more than one AE in the same hospitalization, and in this case the study collects all of them.

6. Statistics – this section describes prevalence and incidence being estimated through logistic regression taking into account covariates. However, the abstract presents crude rates. It should be clearer in the abstract and results which figures are crude and which are adjusted. Confidence intervals should be presented for all estimates in the text and tables.

RESULTS: The prevalence of adverse events was 10,46% (95%CI: 9,91 to 11,04) (1206/11379), while the accumulated incidence in the retrospective incidence study was 28,9% (95%CI:25,9 to 31,2) (314/1088).

Table3. Correlates of adverse events in multiple logistic regression analyses.

	PREVALENCE				INCIDENCE			
Variables	p-value	OR	95% CI for OR		p-value	OR	95% CI for OR	
Department (1)	0,00				0,00			
Surgery and gynaecology	0,06	1,17	0,99	1,38	0,01	1,75	1,17	2,61
Obstetrics	0,02	1,37	1,06	1,78	0,05	0,38	0,15	0,99
Paediatrics	0,00	1,50	1,21	1,85	0,15	0,40	0,12	1,40
Intensive care	0,00	2,52	1,96	3,26	0,01	2,77	1,25	6,17
Complexity of the hospital (2) tertiary	0,02	1,45	1,07	1,97				
Type of admission (3) urgent	0,00	1,34	1,12	1,61	0,59	1,14	0,71	1,83
Length of stay until the day of study	0,03	1,00	1,00	1,00	0,01	1,00	1,00	1,01
Patient comorbidty (4) any	0,00	1,42	1,22	1,64	0,00	2,02	1,28	3,19
Use of medical devices (5) any	0,00	2,59	2,14	3,14	0,00	3,24	1,79	5,85
Country 1	0,00				0,00			
Country *2*	0,00	0,46	0,38	0,56	0,00	0,34	0,20	0,59
Country *3*	0,37	0,91	0,73	1,12	0,16	1,44	0,86	2,41
Country *4*	0,00	0,65	0,52	0,81	0,00	0,22	0,12	0,42
Country *5*	0,04	0,82	0,69	0,99	0,01	0,38	0,19	0,76
Constant	0,00	0,04			0,00	0,03		

Reference categories: (1) medical specialties; (2) secondary hospitals of intermediate complexity with at least surgical theatres, and postsurgical resuscitation wards; (3) planned admission; (4) and (5) absence of risk factors.

7. Results - the figures presented in the abstract should also appear in the text of the results section.
We have added the results in a tables 2 and 3 in the results section.

8. Discussion - the adverse event rate of 28.9% is significantly above all other national/large adverse events studies. However the methods description is incomplete and it is difficult to be certain that the methods used are comparable to other studies. If they are then this is a very significant result and should be further discussed. The current discussion is unclear and lacks any referencing.

The frequency of AE in both the prevalence and incidence studies was greater than that found in previous studies, which may be due to the different characteristics of the patients, who had a higher average age and more risk factors. The nature of the sample selection and the peculiarities of the different search systems prevent statistical inferences and comparisons either within each country or between the countries which are part of the study.

In any case, higher prevalence means higher incidence. In some way the interdependence of these frequency measures remains when we use prevalence on a given day. The fact that prevalence is sensitive to the differences in the characteristics of the patients and that it reflects the differences found between countries, would make it a useful tool in the study and follow-up of the frequency of AE and in comparative studies. Furthermore, as the explanatory model for the occurrence of AE is the same, studying the factors which influence prevalence may provide the same clues when designing strategies for AE control and therefore provide a more efficient tool.

9. Information on resources used in both types of study would be interesting (cost, numbers of reviewers, time taken).

Training and data collection time and cost are not easy to estimate, because they were not calculated.

Specific comments:

1. Abstract – results section: explain what the range refers to. Confidence intervals should be provided.

RESULTS: The prevalence of adverse events was 10,46% (95%CI: 9,91 to 11,04) (1206/11379), while the accumulated incidence in the retrospective incidence study was 28,9% (95%CI:25,9 to 31,2) (314/1088)

2. Summary box – the first sentence seems incomplete 'The identification of adverse events is the first step to improve.'

The identification of adverse events is the first step to improve patient safety

3. Introduction:

– paragraph two has only one citation for the first sentence and its subsequent statements are inadequately referenced.

– paragraph three sentence one should be 'consists of' (not in). Please provide a reference in English for this important statement introducing cross-sectional surveys.

Every research methodology and data collection system has their own caveats[v]. Routine information systems have limitations related to compliance and coding bias. Events reporting systems also show preference for the type of events that reporters consider more relevant and have difficulties tracing duplicates, in addition to still facing unresolved legal issues in many contexts, which penalize reporting and limit their effective use. Prospective studies tend to focus on the analysis of higher-risk patients in detriment of other patients. Medical records, electronic or not, are threatened as well by lack of completeness and recording bias, since clinicians tend to record the data that is more meaningful to them from a clinical point of view. In addition, medical retrospective records review involve a sophisticated sampling plan, and a resource intensive process of record retrieval, reviewing and abstracting of sometimes very heavy and complex medical records.

A data collection process that has been less frequently used in the field of patient safety research, despite its potential, consists of running periodic cross-sectional surveys aiming to assess the point prevalence of AE[vi,vii]. This design has been commonly used to monitor the frequency of healthcare associated infections in many hospitals across Europe and elsewhere[viii], where it has proven to be a feasible and valid methodology, capable to be run with no excessive resources at large scale and across many institutions and organizational cultures. Among the advantages of this design are that instead of requiring a statistically savvy sampling plan, all patients admitted at a given time to the hospital can be surveyed at once, simplifying the sampling process as well as the search and retrieval of records from the archives, since these are usually located near the patients in the wards[ix]. This design also gives researchers the opportunity to ask the attending clinicians for some clarifications in the records, including some missing data. The unit of observation in this design is typically one day of admission, which makes it much shorter and simpler for the reviewers, and gives an estimate of a one-day prevalence, as opposed to the period or incidence rate of a retrospective record review[x].

Because of its greater simplicity, the management of large and multi-centered research studies is also simplified.

[5] Corrales MJ, Limón R, Miralles JJ, Gea MT, Requena J, Aranaz JM, Grupo de trabajo del proyecto EPIDEA. Factores asociados a las infecciones evitables relacionadas con la atención sanitaria identificadas en el estudio EPIDEA. Medicina Preventiva 2010; 16:18-23.

[6] Zarb P, Coignard B, Griskeviciene J, Muller A, Vankerckhoven V, Weist K, Goossens M, Vaerenberg S, Hopkins S, Catry B, Monnet D, Goossens H, Suetens C. The European Centre for Disease Prevention and Control (ECDC) pilot point prevalence survey of healthcare-associated infections and antimicrobial use. Euro Surveill. 2012 Nov 15;17(46)

[7] Sedgwick P.Bias in observational study designs: cross sectional studies. BMJ. 2015 Mar 6;350 :h1286.

[8] Philippe M, Olsen S, Saillour-Glénisson F, Limón R, Aibar C, Aranaz J. Assessing and tackling patient harm: a methodological guide for data-poor hospitals. Geneva: World Health Organization, 2010. Available in:

http://www.who.int/patientsafety/research/methodological_guide/PSP_MethGuid.pdf

[9]Aranaz-Andrés JM, Aibar-Remón C, Limón-Ramírez R, Amarilla A, Restrepo FR, Urroz O, Sarabia O, Inga R, Santivañez A, Gonseth-García J, Larizgoitia-Jauregui I, Agra-Varela Y, Terol-García E. Diseño del estudio IBEAS: prevalencia de efectos adversos en hospitales de Latinoamérica. Rev Calidad Asistencial 2011.

10 Michel P, Quenon JL, De Sarasqueta, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. BMJ 2004; 328(4): 199-203

4. The definition of an adverse event used in IBEAS should be provided and stated whether this was the same for both studies.

The AE definitions used in both designs were those published by WHO in the International Classification for Patient Safety.[xi] Conceptual Framework for the International Classification fo rPatient Safety. Disponible http://www.who.int/patientsafety/taxonomy/icps_full_report.pdf

6. Methods - pg 5, line 25. Provide an exact sample size.
The number of patients included in the cross sectional study was 11.379 patients. 3853 of them (33.9%) fulfilled at least one of the screening criteria. In the second phase of the cross-sectional study 1.191 patients had an AE, which means a prevalence of AE of 10,46% (CI 95%: 9,91 to 11,04).

7. Pg 5, line 27 'An AE was considered prevalent..' should this be present?

We consider it could be helpul to understand the results:

A prevalent AE is defined as one that originates during hospitalization and is clinically present on the day of the study, either as an after-effect or under treatment. This also includes those AE that were occasioned prior to hospitalization at any care level and which led to subsequent admission. AE that had occurred prior to the survey and whose effects had disappeared without prolonging the hospitalization on that particular day were not included.

8. Results – the results of the full IBEAS study should be presented in the text as well as those of the one-day study.

We have referenced this in the bibliography. We also present a ppt with the most relevant results.

[1]2 Aranaz Andres JM, Aibar-Remón C, Limón-Ramírez R, Amarilla A, Restrepo FR, Urroz O et al. Prevalence of adverse events in the hospitals of five Latin American countries: results of the 'Iberoamerican study of adverse events' (IBEAS) BMJ Qual Saf 2011; 20:1043-1051.

9. Pg 7 line 6 – there were 215 patients where the incident was caused by healthcare management but there were 314 adverse events. Does this mean that more than one adverse event was able to be counted per patient? This should be made clearer in the methods and the results.

A patient could have more than one AE in the same hospitalization, and in this case the study collects all of them.

10. Pg 8 line 3 'Surgical patients were associated with more risks than patients admitted in the medical wards' – but the confidence interval for the odds ratio of the one-day study included 1 (OR 1.17 (95% CI 0.99-1.38))?

In both studies, the highest risk of suffering adverse events was seen in ICU patients. Surgical patients were associated with more risks than patients admitted in the medical wards. Whereas, in the cross sectional prevalence study, obstetrics and pediatric patients also showed higher risk than medical patients.
In cross sectional study OR 1,17 (0,99 a 1,38)
In incidence study OR 1,75 (1,17 to 2,61)

11. Pg 8 line 8 'the stay before survey' – what is this? Please define. How long prior to the admission under examination did you look? Also in table 3.

It refers to length of stay until the day of study

12. Pg 8 line 47 – what is 'paradigmatic'?
It means the most frequent

13. Pg 8 last para should refer to Table 4. Because no confidence intervals are given how can one justify the conclusion that retrospective review identifies less serious and shorter adverse events etc? The adverse events identified through the prevalence study seemed to be associated more frequently with hospital readmissions and slightly more with prolonged stay, whereas the frequency of adverse events which did not caused prolonged stay or readmission was higher in the retrospective review (Table 5).

14. Pg 9 line 39 – what does this mean 'Health related infections ….are considered less preventable than other types of AE and these categories are identified further in the prevalence study.'? Please explain. This information may be better suited to the discussion section.
It means Heathcare associated infections (HAI), also called nosocomial infections.

15. Pg 9 line 42 – provide the data or state data not shown and a confidence interval for the statement 'The severity of the AE was not associated with its preventability'.
We have eliminated this comment

16. Discussion – pg 10 the first three paragraphs on study types require referencing.

The choice of the most appropriate epidemiological method in the study of magnitude of AE is not a trivial issue. The question has been analysed in different studies and the generalized consensus is that the choice of method should be based on the aims of the study and the need to combine the minimization of bias and the validity of AE identification with the reproducibility of value judgements on their iatrogenic nature and/or preventability(17[xii].

[1]17 Aranaz-Andrés JM, Aibar-Remón C, Vitaller-Murillo J, Ruiz-López P, Limón-Ramírez R, Terol-García E; ENEAS work group. Incidence of adverse events related to health care in Spain: results of the Spanish National Study of Adverse Events. J Epidemiol Community Health. 2008 Dec; 62(12):1022-9.

17. Pg 10, line 25 Please explain or reference the statement 'The relationship between prevalence and incidence generally depends on the duration of the event under review and the period of observation.' This would not be the standard definition for these terms nor how they have been used in previous adverse event studies.
Prevance in epidemiology is related to the disease duration. For long processes, with less mortality for example. prevalence is higher

18. Figures 1a and 1b – label arrows. The figure notes contain definitions which should also be included in the methods.
We have included this terms in methods section

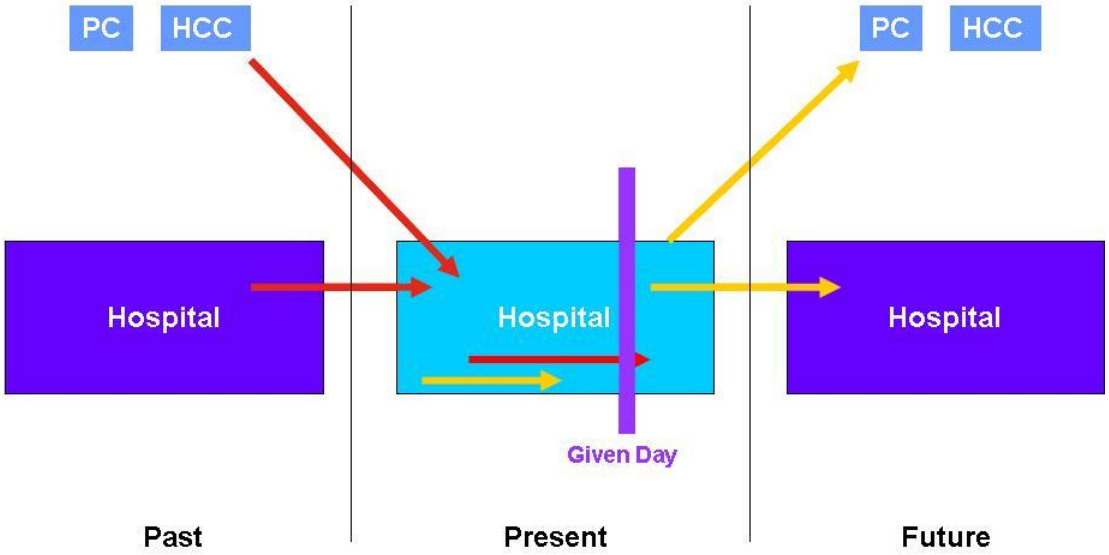

Figure note: For this study, a prevalent AE is defined as one that originates during hospitalization and is clinically present on the day of the study, either as an after-effect or under treatment. This also includes those AE that were occasioned prior to hospitalization at any care level and which led to subsequent admission.

Figure 1b. Scheme of incident AE study. PC: Primary Care, HCC: Healthcare Centre

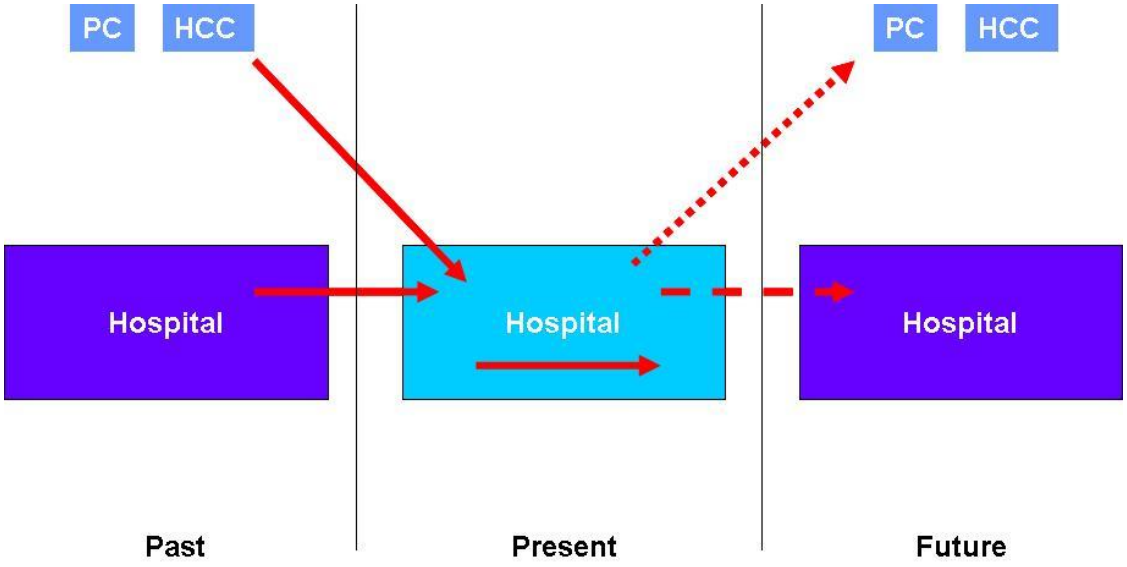

Figure note: An AE incident is defined as one that occurs during any patient care process, as it may be detected at another level of care or in other hospital. In practical terms, as we carried out a retrospective study based on clinical hospital records, primary care AE were not included. Those that led to readmission in the same or another hospital were compensated by the AE which were detected during this hospitalization and which had been originated in a previous hospitalization.

19. Pg 11 line 31 – what do you mean by 'we will be able to detect the number of false positive through the revision of the second questionnaire (MRF2)' and 'This evaluation will enable us to adjust the alarm conditions for transversal studies'? Please explain these statements further.

In both studies (cross sectional and retrospective), researchers used two tools to detect harmful incidents, namely a Screening Guide and a Modular Questionnaire[xiii,xiv] to identify harmful incidents using the medical record review methodology[xv,xvi].

First, the screening guide was applied to the patients in the study. This served as an alert and tracking system for possible incidents. All the patients admitted at hospital (except emergency room) were studied. The screening was made by well trained nurses.

If a patient screened positive for one or more of the 19 alert criteria in the screening guide, the case was studied using the case history. An in-depth study of case histories enabled researchers to conclude whether a patient did in fact present with the consequences of a harmful incident (true positive) and if so, to classify the type of event, its severity, any associated factors, and whether or not the incident could have been avoided, etc. This second confirmatory review was made (in both cross sectional and retrospective) by medical doctors with at least 5 years of clinical experience. A patient could have more than one AE in the same hospitalization, and in this case the study collects all of them.

20. Pg 11 line 41 - provide the references for the other AE studies referred.

Australia:

Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study.Med J Aust. 1995;163:458-71.

Nueva Zelanda:
Brow P, McArthur C, Newby L et al. Cost of medical injury in New Zealand: a retrospective cohort study. J Health Serv Res Policy 2002;7:29–34.

United Kingdom:
Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. BMJ.2001;322:517-9
Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice. J R Soc Med. 2001; 94:322-30.

Canadá:
Forster AJ, Asmis TR, Clark HD, Saied GA, Code CC, Caughey SC, et al. Ottawa Hospital Patient Safety Study: incidence and timing of adverse events in patients admitted to a canadian teaching hospital. Can Med Assoc. 2004;170:1235-40.

USA:
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.
Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med* 1991;324(6):377-84.
Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000;38(3):261-71.

21. Pg 11 last para – this should be in the methods. Nosological means classification of disease, is this the term you wish to use?
We have changed to pathological process

22. Pg 12 first sentence ‘the reliability of the questionnaire in other studies has been assessed as moderate’ – please reference.
Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice. J R Soc Med. 2001; 94:322-30.

We have added all the corrections for the tables.
Tables
– Table 1 - why was a p level of 0.1 chosen to indicate the level of non significance in table 1?
– Confidence intervals should be provided.
– The columns are not fully labeled.
– The studies should be presented consistently (eg one the left and the other on the right - table 1 differs from the others.
– Table 5 – are these crude or adjusted figures? What is SD and ST?

Reviewer: 2
Reviewer Name: Dr Sharon Mayor
Institution and Country: Division of Population Medicine, Cardiff School of Medicine, Cardiff, Wales UK Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below This is a well-conceived study addressing an important question regarding the optimal methodological approach in undertaken case record review to improve the quality of care provided. Interesting data is generated, however, I think it would be useful to see greater clarity in this paper on a number of methodological points and a little more clinical application and insight in the discussion.

Methods

Firstly, in describing the methods, it would be helpful to have a brief overview on the training provided to the reviewers.
Training of reviewers:
The reviewers training took place in two stages. First, the trainer workshop addressed the national coordinating teams. Second, the national coordinators trained in turn the national investigators. A concordance study was carried out using clinical records from each country.

Secondly, previous studies report a modest level of inter rater reliability using the structured review forms used and if this was not assessed in this study, it should be raised as a limitation as it might explain some of the inter-country variation seen in the reported rates of AEs.

Table 1 Kappa index for screening, identification and preventability of adverse events

	Positive screening	Identified adverse events	Preventable adverse events
Country 1			0.62
Country 2	0.85	0.87	0.74
Country 3		0.32	
Country 4		0.38	0.47
Country 5	0.55	0.30	0.27
Other studies	0.70	0.40–0.80	0.19–0.69

Thirdly, it would be helpful to state up front in the methods that a Likert scale was used to determine the presence and preventability of detected AEs and that the threshold of 4/6 was used in both case?

We send the MRF2

Fourthly, the authors state the second phase of the study took place most frequently after the patient's discharge. Can the authors clarify this proportion as the incidence study seeks to examine the whole of the inpatient episode and if a significant proportion of reviews were undertaken whilst the patient was still an inpatient the rates might be underestimated?
The screening phase of the retrospective review found about 44,5% of the medical records, corresponding to 484 patients, positive for at least one of the 19 triggers included in the forms. At the confirmatory phase, it was determined that 40 of those patients had experienced one or more incidents without harm or prolonged stay, and 288 patients had experienced at least one harmful patient safety incident. Of these, in 215 patients the incident was considered to be mostly related to the health care received rather than to the patient intrinsic vulnerability. Thus the proportion of patients suffering at least an adverse event related to the care received before or during their hospitalization was close to 19,8% (95% CI: 17,2 to 21,9). In total, there were 314 AE (because a patient could have more than one AE) related to healthcare corresponding to a retrospective incidence of AE 28,9% (95% CI: 25,9 to 31,2).

Results

On page 7 the authors state the proportion of patients experiencing an adverse event was close to 19.8%. However, they then report a total of 314 adverse events and a retrospective incidence of

28.9%. Is this as a result of multiple adverse events in individual cases or has a lower threshold been used in determining causality? Can this be clarified?

Yes, a patient can suffer several AE.

The interpretation of the different rates and types of events in both studies is interesting but does highlight a number of important issues if the findings are going to be used to prioritise improvement efforts. For example, nosocomial pneumonia and neonatal complications were found to be serious AEs in the prevalence study with nosocomial pneumonia probably being of lower preventability and neonatal complications being infrequent. However, the less serious issues identified through reviewing the whole episode of care e.g UTIs and wound infections, may be more useful in identifying where improvement priorities should be focused.

Discussion

Can the authors be sure that the reviewers were trained enough to perform the review in a reliable way.
The reviewers training took place in two stages. First, the trainer workshop addressed the national coordinating teams. Second, the national coordinators trained in turn the national investigators. A concordance study was carried out using clinical records from each country

Some countries reported rates in line with previous studies, others were three times higher than what would be expected in other settings.

We presente the heterogeneity seen between hospitals in the 5 different countries

Table 2 Characteristics of the study population

	Country 1	Country 2	Country 3	Country 4	Country 5	Total
Patients studied	2373 (20.9%)	2897 (25.5%)	1632 (14.3%)	2003 (17.6%)	2474 (21.7%)	11 379 (100.0%)
Median age	49 (44)	35 (33)	50 (36)	42 (40)	55 (40)	45 (39)
Women	1129 (47.6%)	1544 (53.3%)	855 (52.4%)	1161 (58.0%)	1286 (52.0%)	5975 (52.5%)
Specialised medical care	887 (37.4%)	751 (25.9%)	658 (40.3%)	710 (35.4%)	1039 (42.0%)	4045 (35.5%)
Surgery/gynaecology	739 (31.1%)	1099 (37.9%)	509 (31.2%)	762 (38.0%)	789 (31.9%)	3898 (34.3%)
Obstetrics	142 (6.0%)	362 (12.5%)	231 (14.2%)	305 (15.2%)	201 (8.1%)	1241 (10.9%)
Paediatrics	453 (19.1%)	561 (19.4%)	168 (10.3%)	205 (10.2%)	314 (12.7%)	1701 (14.9%)
Intensive care	152 (6.4%)	124 (4.3%)	66 (4.0%)	21 (1.0%)	131 (5.3%)	494 (4.3%)
Emergency admission	1523 (86.5%)	2189 (81.2%)	1291 (84.3%)	1153 (67.2%)	1875 (77.3%)	8031 (79.3%)
Planned admission	238 (13.5%)	506 (18.8%)	241 (15.7%)	564 (32.8%)	550 (22.7%)	2099 (20.7%)
Stay before the survey	5 (10)	7 (13)	6 (14)	3 (6)	8 (17)	6 (13)
Commorbidity factors	1238 (52.2%)	1507 (52.00%)	854 (52.30%)	1103 (55.10%)	1426 (57.60%)	6128 (53.90%)
Risk factor associated with the use of medical devices	1935 (81.50%)	2472 (85.30%)	884 (54.20%)	1633 (81.50%)	1560 (63.10%)	8484 (74.60%)

Serious adverse events are highlighted as being important but these are relatively infrequent. Other AEs such as skin and pressure damage and patient falls etc are unlikely to be detected in one-day prevalence surveys but offer important insights into the general quality of care provided?
Yes we agree at this poity. It is a limitation of the study.

Should the discussion include more analysis of the relative strengths and limitations of the two designs when used for different purposes.

The fact that prevalence is sensitive to the differences in the characteristics of the patients and that it reflects the differences found between countries, would make it a useful tool in the study and follow-up of the frequency of AE and in comparative studies. Furthermore, as the explanatory model for the occurrence of AE is the same, studying the factors which influence prevalence may provide the same clues when designing strategies for AE control and therefore provide a more efficient tool.

Moreover, the fact that the prevalence design detects proportionally more serious AE is not a drawback. On the contrary, these are precisely the AE which need to be prioritized when designing control strategies, and as we commented above, the detected AE were equally preventable in both

designs. This reinforces the idea that preventability and seriousness of the EA are independent factors.

As the point prevalence design is more efficient in terms of time and resources, its validity is less dependent on the quality of the clinical records and allows simultaneous study through other observation and audit systems, regular prevalence on a given day studies might provide an efficient AE monitoring and control strategy.

VERSION 3 – REVIEW

REVIEWER	Natasha Rafter Royal College of Surgeons in Ireland Ireland
REVIEW RETURNED	20-Apr-2017

GENERAL COMMENTS	Thank you to the authors for their responses to my initial review. However, I still have a number of concerns that need to be addressed. Although the distinction between prevalent and incident AEs is defined, a definition of an adverse event itself is not present in the paper. Somewhat confusingly the term 'harmful incidents' is also used and is not defined. It is thus not clear what constitutes either a harmful incident or an adverse event and how these terms relate to one another. I note that an adverse event was defined in the IBEAS BMJQ&S paper ('An AE was defined as any event causing harm to the patient that was perceived to be more related to the healthcare management rather than to the patient's underlying condition'); if applicable to both studies in the current paper this should be included in the methods also.[1] Furthermore in the results section there is reference to incidents then patient safety incidents then adverse events, however this process of moving from incidents to patient safety incidents to adverse events is not explained nor are the terms defined in the methods. The criteria for determining preventability are still not included in the paper. Instead readers are directed to the Modular questionnaire which is referenced. Given that the level of preventability is quoted in the results the reader should be able to assess its meaning from the paper. Therefore I recommend the criteria for determining preventability are described either in the paper or in an appendix. The headline results are not consistent between the abstract and the results. For example, the abstract refers to a prevalence of 10.46% calculated as 1206 patients with an AE out of 11379 (1206/11379). However the results section states 1191 patients had an AE. Please correct this discrepancy. The paper reports the retrospective study screening and review process in detail under Table 1. This description would be enhanced by the use of a flowchart. The corresponding cross sectional study screening and review process should be presented here also in order for the current paper to be understandable (and the methods of the two studies able to be compared) without needing to refer to a separate publication.

	The prevalence/incidence nomenclature is a problem for this field of research and has been used differently in other record review studies. Hence I am uncomfortable with the use of the terms 'prevalence study' and 'incidence study' and would prefer they were described by their method (cross sectional or one day study and retrospective record review study). This would be consistent with the Michel et al paper which compared three methods of determining adverse events.[2] Although the authors state they agree with this comment and state that they have changed the terms in their response to my comments, the text still refers to prevalence and incidence studies as do tables 3-6. The calculation of the adverse event rate for each study is not well described in the methods – in particular which figures are used for the numerator and denominator. It appears that the two studies are being compared with different adverse event rates – the cross sectional study uses number of patients with an AE in the numerator (1191 patients with AEs out of 11379 patients is 10.46%) whilst the retrospective record review study uses number of AEs in the numerator (314 AEs in 1088 patients is 28.9%, note that this is described as an 'accumulated incidence' in the abstract and in table 2 - do the authors mean incidence density?). The paper is therefore using two different statistics to compare the two studies when the comparison of adverse event estimation in this paper should use the same statistic (i.e. with the same numerator). Thus the retrospective record review figure of 19.8% (215 patients with AEs out of 1088), not the 28.9%, should be the one used to compare with the cross sectional study result of 10.46%. This comparison would be consistent with the method of Michel et al and with the main results published by the current first author in the Spanish retrospective record review study.[2, 3] In addition, 19.8% is also the more suitable adverse event occurrence rate to use for comparisons with other retrospective record review studies in the field.[3, 4] This paper would benefit from further review of its readability and use of technical terms, for example accumulated incidence and transversal are not standard terms. Confidence intervals should be presented around the adverse event estimates in tables 2, 4, and 5. In addition, the authors should test formally for discrepancy between the studies with p values. 1. Aranaz-Andres, J.M., et al., Prevalence of adverse events in the hospitals of five Latin American countries: results of the 'Iberoamerican study of adverse events' (IBEAS). BMJ Quality & Safety, 2011. 2. Michel, P., et al., Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. British Medical Journal, 2004. 328: p. 199-204. 3. Aranaz-Andres, J.M., et al., Incidence of adverse events related to health care in Spain: results of the Spanish national study of adverse events. Journal of Epidemiology and Community Health, 2008. 62: p. 1022-9. 4. de Vries, E.N., et al., The incidence and nature of in-hospital adverse events: a systematic review. Quality Safety Health Care, 2008. 17: p. 216-223.

VERSION 3 – AUTHOR RESPONSE

1. Although the distinction between prevalent and incident AEs is defined, a definition of an adverse event itself is not present in the paper. Somewhat confusingly the term 'harmful incidents' is also used and is not defined. It is thus not clear what constitutes either a harmful incident or an adverse event and how these terms relate to one another. I note that an adverse event was defined in the IBEAS BMJQ&S paper ('An AE was defined as any event causing harm to the patient that was perceived to be more related to the healthcare management rather than to the patient's underlying condition'); if applicable to both studies in the current paper this should be included in the methods also.

Response: The Conceptual Framework for the International Classification for Patient Safety defines a harmful incident as a synonym of adverse event:

A harmful incident (adverse event) is an incident that results in harm to a patient (e.g., the wrong unit of blood was infused and the patient died from a haemolytic reaction).

Harm implies impairment of structure or function of the body and/or any deleterious effect arising there from, including disease, injury, suffering, disability and death, and may be physical, social or psychological.

We have included this definition in the text in the last introduction paragraph, and we have added also the definition of adverse event used in the IBEAS BMJ Q&S paper.

2. Furthermore in the results section there is reference to incidents then patient safety incidents then adverse events, however this process of moving from incidents to patient safety incidents to adverse events is not explained nor are the terms defined in the methods.

Response: The Conceptual Framework for the International Classification for Patient Safety defines:

A patient safety incident is an event or circumstance that could have resulted, or did result, in unnecessary harm to a patient. In the context of the ICPS, a patient safety incident will be referred to an incident.

However, to maintain a similar terminology we have changed all the terms in the results, and we have chosen the term "patient safety incidents" for "patient safety incidents without harm", and "adverse events" referring to "harmful patient safety incident".

3. The criteria for determining preventability are still not included in the paper. Instead readers are directed to the Modular questionnaire which is referenced. Given that the level of preventability is quoted in the results the reader should be able to assess its meaning from the paper.
Therefore I recommend the criteria for determining preventability are described either in the paper or in an appendix.

Response: Preventable is being accepted by the community as avoidable in the particular set of circumstances.
Preventability: To mine the preventability of the AE's, the possibility of their being prevented was scored on a 1-6 scale (1 = no evidence of preventability; 6= total evidence).
Those AE's score within the 1-3 range were considered unpreventable or hardly preventable, those scoring higher than 3 on this scale being considered preventable.
1. No evidence of preventable AE

2. Minimal probability of preventable AE
3. Slight probability of preventable AE
4. Moderate probability of preventable AE
5. Highly probable of preventable AE
6. Total evidence of preventable AE

This definition is proposed in Modular Questionnaire MRF2, in:
T.A. Brennan,L.L. Leape,N.M. Laird,L. Hebert,A.R. Localio,A.G. Lawthers
Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med, 324 (1991), pp. 370-376

We have added this in appendix 1, as the reviewer suggested.

4. The headline results are not consistent between the abstract and the results. For example, the abstract refers to a prevalence of 10.46% calculated as 1206 patients with an AE out of 11379 (1206/11379). However the results section states 1191 patients had an AE. Please correct this discrepancy.

Response: We have clarified this in the abstract, referring as it is suggested as prevalence of patients with AE of 10,47% (1191/11379) in the cross-sectional study, and cumulative incidence of patients with AE of 19,76% (215/1088), in the retrospective cohort study.

As a patient can suffer more than one AE in the hospitalization, we have also added the total number of AE detected. In the cross-sectional study the prevalence of total AE was 11,85% (1349/11379), and in the retrospective cohort study the cumulative incidence ot total AE was 28,86% (314/1088).

There was an error in the previous document, the total number of AE detected in the cross-sectional study was 1349, not 1206.

We have added Table 2 with this result to clarify them.

5. The paper reports the retrospective study screening and review process in detail under Table 1. This description would be enhanced by the use of a flowchart. The corresponding cross sectional study screening and review process should be presented here also in order for the current paper to be understandable (and the methods of the two studies able to be compared) without needing to refer to a separate publication.

Response: We have added these flowcharts for cross-sectional study and for the retrospective one, respectively.

Flow chart 1: Patients studied in the cross sectional study.
Flow chart 2: Patients studied in the retrospective incidence study.

6. The prevalence/incidence nomenclature is a problem for this field of research and has been used differently in other record review studies. Hence I am uncomfortable with the use of the terms ‘prevalence study’ and ‘incidence study’ and would prefer they were described by their method (cross sectional or one day study and retrospective record review study). This would be consistent with the

Michel et al paper which compared three methods of determining adverse events.[2] Although the authors state they agree with this comment and state that they have changed the terms in their response to my comments, the text still refers to prevalence and incidence studies as do tables 3-6.

Response: We have modified the terms. The paper now refers to cross-sectional and retrospective cohort study.

Some definitions of the Dictionary of Epidemiology could improve to clarify this aspects:

-cross sectional study and prevalence study are synonymous
-cumulative incidence is also the incidence proportion
-retropective study: our study is a cohort study historical, that means, prospective study in retrospect.

Reference: A dictionary of epidemiology. Edited for the International Epidemiological Association by Miquel Porta. Sixth Edition. 2014. Oxford University Press. ISBN 178-0-19-997673

7. The calculation of the adverse event rate for each study is not well described in the methods – in particular which figures are used for the numerator and denominator. It appears that the two studies are being compared with different adverse event rates – the cross sectional study uses number of patients with an AE in the numerator (1191 patients with AEs out of 11379 patients is 10.46%) whilst the retrospective record review study uses number of AEs in the numerator (314 AEs in 1088 patients is 28.9%, note that this is described as an 'accumulated incidence' in the abstract and in table 2 - do the authors mean incidence density?). The paper is therefore using two different statistics to compare the two studies when the comparison of adverse event estimation in this paper should use the same statistic (i.e. with the same numerator). Thus the retrospective record review figure of 19.8% (215 patients with AEs out of 1088), not the 28.9%, should be the one used to compare with the cross sectional study result of 10.46%. This comparison would be consistent with the method of Michel et al and with the main results published by the current first author in the Spanish retrospective record review study.[2, 3] In addition, 19.8% is also the more suitable adverse event occurrence rate to use for comparisons with other retrospective record review studies in the field.

There are two main result measures:

1) Proportion or percentage of patients with adverse events:

a. Numerator: number of patients with at least one adverse events (or patient safety incident with harm, related to the healthcare) multiplied by 100

b. Denominator: number of studied patients.

2) Proportion or percentage of adverse events:

a. Numerator: total number of adverse events detected (because a patient could have more than one AE) multiplied by 100

b. Denominator: number of studied patients.

In the cross sectional design we use prevalence term.
In the retrospective design we use the cumulative incidence or incidence proportion terms.

Cross sectional design:
• Prevalence (proportion or percentage) of patients with adverse events:

o 1191/11379=10,47% (CI 95%: 9,90 to 11,03)
o Numerator: number of patients with at least one adverse events (or patient safety incident with harm, related to the healthcare) multiplied by 100
o Denominator: number of studied patients.

• Prevalence (proportion or percentage) of adverse events:

o 1349/11379= 11,85% (CI 95%: 11,26 to 12,46)
o Numerator: total number of adverse events detected (because a patient could have more than one AE) multiplied by 100
o Denominator: number of studied patients

Retrospective design

• Cumulative incidence (or incidence proportion) of patients with adverse events:

o 215/1088=19,76% (CI 95%: 17,35 to 22,17)
o Numerator: number of patients with at least one adverse events (or patient safety incident with harm, related to the healthcare) multiplied by 100
o Denominator: number of studied patients.

• Cumulative incidence (or incidence proportion) of adverse events:

o 314/1088= 28,86% (CI 95%: 26,12 to 31,6)
o Numerator: total number of adverse events detected (because a patient could have more than one AE) multiplied by 100
o Denominator: number of studied patients

We have added Table 2 with this main results to clarify them.

Cross sectional
(prevalence) Retrospective
(cumulative incidence)
Patients with AE 1191/11379=10,47%
(CI 95%: 9,90 to 11,03) 215/1088=19,76%
(CI 95%: 17,35 to 22,17)

Total Lumber of AE 1349/11379= 11,85%
(CI 95%: 11,26 to 12,46) 314/1088= 28,86%
(CI 95%: 26,12 to 31,6)

7. This paper would benefit from further review of its readability and use of technical terms, for example accumulated incidence and transversal are not standard terms.

Response: We have modified this. We use cumulative incidence, and tranversal is no more used in the paper.

8. Confidence intervals should be presented around the adverse event estimates in tables 2, 4, and 5.

Response: We have added the confidence intervals in tables 2, 4, 5.

9. In addition, the authors should test formally for discrepancy between the studies with p values.

Response: We have preferred not to test with a chi2 the main results of this study because we consider that a cross-sectional study and an retrospective cohort one cannot be tested in an statistical way, because they are obtained with two different study designs.

VERSION 4 – REVIEW

REVIEWER	Natasha Rafter Royal College of Surgeons in Ireland Ireland
REVIEW RETURNED	07-Aug-2017

GENERAL COMMENTS	The authors have addressed my previous comments.

VERSION 4 – AUTHOR RESPONSE

Dear all

We are grateful in sending you this manuscript for evaluation.

We have added this paragraph in the dicussion section:

A limitation of this study could be the sample used to evaluate the retrospective cohort. It is a 10% of medical records used in the cross-sectional study and proportional to hospital size, which could be not representative of all the population attended in the health system. Tertiary hospitals have more complexity and it could be overestimated the number of adverse events, in comparison with the total number of patients attended in the country. Another limitation is the quality of medical records. If the variability in the accomplishment between the different countries were high the comparison between them could be weaker.

We also have clarified the competing interests in the Article Summary.
We have added the strengths and limitations in the Article Summary.

---

[i] O'Neil AC, Petersen LA, Cook EF, Bates DB, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. Ann Intern Med 1993;119:370-6.

[ii] Michel P, Quenon JL, Sarasqueta AM, Scemama O. L'estimation du risque iatrogène graves dans les établissements de santé en France: les enseignements d'une étude pilote dans the région Aquitaine. Etudes et Résultats 2003; 219:1-8.

[iii] Michel P, Aranaz JM, Limón R, Requena J. Siguiendo the pista de los efectos adversos: Cómo detectarlos. Rev Calidad Asistencial 2005;20:204-10.

[iv] Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgments concerning adverse events suffered by hospitalized patients. Med care;27(12):1148-58

[v] Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med. 2003 Jan;18(1):61-7

[vi] Requena J, Aranaz JM, Gea MT, Limón R, Miralles JJ, Vitaller J, Grupo de trabajo del proyecto EPIDE. Evolución de la prevalencia de efectos adversos relacionados con la asistencia en hospitales de la Comunidad Valenciana Rev Calidad Asistencial 2010; 25:244-9.

[vii] Corrales MJ, Limón R, Miralles JJ, Gea MT, Requena J, Aranaz JM, Grupo de trabajo del proyecto EPIDEA. Factores asociados a las infecciones evitables relacionadas con la atención sanitaria identificadas en el estudio EPIDEA. Medicina Preventiva 2010; 16:18-23.

[viii] Zarb P, Coignard B, Griskeviciene J, Muller A, Vankerckhoven V, Weist K, Goossens M, Vaerenberg S, Hopkins S, Catry B, Monnet D, Goossens H, Suetens C. The European Centre for Disease Prevention and Control (ECDC) pilot point prevalence survey of healthcare-associated infections and antimicrobial use. Euro Surveill. 2012 Nov 15;17(46)

[ix] Sedgwick P.Bias in observational study designs: cross sectional studies. BMJ. 2015 Mar 6;350 :h1286.

[x] Philippe M, Olsen S, Saillour-Glénisson F, Limón R, Aibar C, Aranaz J. Assessing and tackling patient harm: a methodological guide for data-poor hospitals. Geneva: World Health Organization, 2010. Available in:

http://www.who.int/patientsafety/research/methodological_guide/PSP_MethGuid.pdf

[xi] Conceptual framework for the International Classification for Patient Safety. Available in http://www.who.int/patientsafety/taxonomy/icps_full_report.pdf

[xii] Aranaz-Andrés JM, Aibar-Remón C, Vitaller-Murillo J, Ruiz-López P, Limón-Ramírez R, Terol-García E; ENEAS work group. Incidence of adverse events related to health care in Spain: results of the Spanish National Study of Adverse Events. J Epidemiol Community Health. 2008 Dec; 62(12):1022-9.

[xiii] O'Neil AC, Petersen LA, Cook EF, Bates DB, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. Ann Intern Med 1993;119:370-6.

[xiv] Michel P, Quenon JL, Sarasqueta AM, Scemama O. L'estimation du risque iatrogène graves dans les établissements de santé en France: les enseignements d'une étude pilote dans the région Aquitaine. Etudes et Résultats 2003; 219:1-8.

[xv] Michel P, Aranaz JM, Limón R, Requena J. Siguiendo the pista de los efectos adversos: Cómo detectarlos. Rev Calidad Asistencial 2005;20:204-10.

[xvi] Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgments concerning adverse events suffered by hospitalized patients. Med care;27(12):1148-58