BMJ

BMJ 2014;348:g3555 doi: 10.1136/bmj.g3555 (Published 4 June 2014)

RESEARCH

Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study

Emma Maund *PhD student*[1], Britta Tendal *postdoctoral researcher*[1], Asbjørn Hróbjartsson *senior researcher*[1], Andreas Lundh *physician*[1 2], Peter C Gøtzsche *professor*[1]

[1]Nordic Cochrane Centre, Rigshospitalet Department 7811, Copenhagen, Denmark ; [2]Department of Infectious Diseases, Hvidovre University Hospital, Hvidovre, Denmark

Abstract

Objective To assess the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports.

Design Systematic electronic search for adverse events of suicidality in tables, narratives, and listings of adverse events in individual patients within clinical study reports. Where possible, for each event we extracted the original term reported by the investigator, the term as coded by the medical coding dictionary, medical coding dictionary used, and the patient's trial identification number. Using the patient's trial identification number, we attempted to reconcile data on the same event between the different formats for presenting data on adverse events within the clinical study report.

Setting 9 randomised placebo controlled trials of duloxetine for major depressive disorder submitted to the European Medicines Agency for marketing approval.

Data sources Clinical study reports obtained from the EMA in 2011.

Results Six trials used the medical coding dictionary COSTART (Coding Symbols for a Thesaurus of Adverse Reaction Terms) and three used MedDRA (Medical Dictionary for Regulatory Activities). Suicides were clearly identifiable in all formats of adverse event data in clinical study reports. Suicide attempts presented in tables included both definitive and provisional diagnoses. Suicidal ideation and preparatory behaviour were obscured in some tables owing to the lack of specificity of the medical coding dictionary, especially COSTART. Furthermore, we found one event of suicidal ideation described in narrative text that was absent from tables and adverse event listings of individual patients. The reason for this is unclear, but may be due to the coding conventions used.

Conclusion Data on adverse events in tables in clinical study reports may not accurately represent the underlying patient data because of the medical dictionaries and coding conventions used. In clinical study reports, the listings of adverse events for individual patients and narratives of adverse events can provide additional information, including original investigator reported adverse event terms, which can enable a more accurate estimate of harms.

Introduction

A proper assessment of the benefits and harms of a medical intervention requires accurate data on harms. An assessment of the harms of an intervention in a randomised clinical trial is more difficult than an assessment of the benefits, as harms can be unpredictable and harms events may be rare.

In a classic drug trial, run and financed by the producer of the drug (the sponsor), doctors interacting with patients (the trial investigators) describe in case report forms those adverse events occurring in each patient, and the sponsor then codes them and enters them in clinical safety databases. The coded data are used for production of summaries of product characteristics and clinical study reports. Clinical study reports comprise detailed information on efficacy and adverse events data from a single trial and can be hundreds of pages in length. These clinical study reports form part of the marketing authorisation application submitted to regulatory authorities, and they should also be used as the primary data source for systematic reviews of drugs.[1] This has been most aptly illustrated by the Cochrane review of neuraminidase inhibitors for preventing and treating influenza in healthy adults and children, where the review based on clinical study reports on oseltamivir (Tamiflu) gave much more modest results than those based on published reports.[2]

In a clinical study report, data on adverse events are presented in various summaries and tabulations, including listings of all

Correspondence to: E Maund em@cochrane.dk

Extra material supplied by the author (see http://www.bmj.com/content/348/bmj.g3555?tab=related#webextra)

Supplementary table 1: adverse events of suicidal ideation and preparatory behaviour
Appendices 1a-c: examples of data formats

adverse events and pre-existing medical conditions in individual patients; narratives of clinically important adverse events (including serious adverse events or discontinuations of the study drug as a result of adverse events), which also include data on pre-existing medical conditions; and summary tables of treatment emergent adverse events (events that occurred or worsened after the study drug was started) or adverse events that emerged after discontinuation of the study drug (see box and supplementary appendices 1a-c for examples of each format).[3]

There are important differences between these three data formats that are related to the coding procedures. The narratives—and in some cases also the listings of data on individual patients—contain the investigator's description of the adverse event on the case report form (commonly referred to as the "verbatim" description). In summary tables, the events appear as coded terms. This is necessary to analyse rates of occurrence because investigators may use different terms to describe the same type of events. The grouping of similar events is achieved by coding verbatim terms to the most closely matching lowest level term in a hierarchically structured medical coding dictionary (tables 1⇓ and 2⇓). Similar lowest level terms are aggregated at the next level into a preferred term, so named because it is a favoured term for use in submissions to regulatory authorities, which are presented in summary tables of adverse events.[10]

Historically, the most widely used dictionaries have been the US Food and Drug Administration's Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART) and the World Health Organization Adverse Reaction Terminology (WHO-ART). These dictionaries were introduced in 1969 in response to increased regulation of the pharmaceutical industry after the thalidomide scandal.[11] However, they had limitations, including lack of specificity of lowest level terms, and in 1999 the International Conference on Harmonisation (ICH) launched the Medical Dictionary for Regulatory Activities (MedDRA).[12] MedDRA has more terms and they are more specific than those in the earlier dictionaries (see tables 1 and 2). For example, a study found that MedDRA contained exact or acceptable matches for 90% of verbatim terms but that COSTART contained only 62%.[13]

MedDRA cannot, however, solve all problems. Firstly, data in summary tables and in listings of adverse events for individual patients may differ from those presented in narratives because of the coding conventions used. For example, the preferred coding convention for a definitive diagnosis with symptoms, such as "anaphylactic reaction, rash, dyspnea, hypotension, and laryngospasm" is to code the diagnosis only as, for example, anaphylactic reaction.[14] In contrast, the preferred coding convention for a provisional diagnosis with symptoms, such as "Possible myocardial infarction with chest pain, dyspnea, diaphoresis" is to code the provisional diagnosis and symptoms as, for example, myocardial infarction, chest pain, dyspnoea, diaphoresis.[14] Secondly, coding can be inconsistent. For example, when the FDA wanted to analyse the risk of suicidality (ideation, behaviour, suicide attempts, and suicide) in paediatric trials of selective serotonin reuptake inhibitors (SSRIs), they found instances of suicidality events coded to both more severe terms and less severe terms. The FDA found that any conclusion based on such data would be unreliable and might lead to either an unwarranted restriction of the drugs or an underestimation of their dangers.[15] Unsurprisingly, research in other areas has shown that misclassifying or omitting even one adverse event can mean the difference between a statistically significant and non-statistically significant association with a drug.[16 17]

We assessed the effects of coding and coding conventions on adverse events data within clinical study reports and compared three different data formats. We used the nine main placebo controlled trials submitted to the European Medicines Agency in the marketing authorisation application of duloxetine for the treatment of major depressive disorder in adults.[18]

Methods

The nine clinical study reports on duloxetine date from September 2000 to September 2003 and total 13 729 pages. We obtained these documents in May 2011 as part of a wider request of access to reports on SSRIs and serotonin norepinephrine reuptake inhibitors. Duloxetine was the only centrally approved product (whereby a single application to the EMA can lead to a European Union wide marketing authorisation for a drug),[19] which is why we focused on this drug. We specifically chose to assess the coding of adverse events of suicidality (ideation, behaviour, attempts, and suicide) within these reports, given the FDA's findings of inconsistency in coding of suicidality in trials of SSRIs in young people, and ongoing public concern and scientific debate about suicidality in adults.[15 20]

One researcher used optical character recognition software to make searchable the 47 PDF documents, comprising the nine clinical study reports. Adobe Acrobat was used for all text portions. ABBYY Finereader enabled the efficient conversion of tables of harms into Excel spreadsheets; according to its manufacturer this software has an accuracy rate of 99.8%.[21]

Two observers, one an experienced medical coder (EM), independently did electronic searches in the clinical study reports of summary tables of coded adverse events, narratives of serious adverse events, narratives of discontinuations of the study drug as a result of adverse events, and listings of adverse events in individual patients. The box describes each data format, and table 3⇓ provides key features of narratives and of individual patient adverse event listings. (See supplementary appendices 1a-c for examples of each data format.)

Search terms included those that the FDA requested pharmaceutical companies to use when searching company databases for events of suicidality in paediatric trials ("suic", "overdos, "attempt", "cut", "gas", "hang", "hung", "jump", "mutilate", "overdos", "self damage", "self harm", "self inflict", "self injur", "shoot", and "slash").[15 22] We additionally used the terms "poi", "emot", "labi", "hos", "vio", "agg", "thought", and "think". We were only interested in adverse events that met the definition in the international statistical principles for clinical trials guideline of a treatment emergent adverse event—that is, an event that occurred or worsened after the study drug was started,[23] or adverse events that emerged after discontinuation of the study drug.

In an Excel spreadsheet we recorded the results of the searches, including which data format the term was found in, which study arm (investigational drug, active comparator, or placebo) the suicidality event occurred in, whether the term found was a verbatim or coded term, and the medical coding dictionary used in the trial. We then compared the extracted data and resolved discrepancies by consensus.

When a verbatim term was reported, one researcher (EM) consulted the medical coding dictionary used in the study and chose the closest matching lowest level term, and then the preferred term that was used in summary tables. As the lowest level terms were not available in the reports, we could not compare our choices with these, but we checked the preferred terms. We accessed COSTART version 5, which was released in 1995 and was the last version of COSTART, through the

Glossary of clinical study report related terms

- Clinical study report (CSR): "A written description of a trial/study of any therapeutic, prophylactic, or diagnostic agent conducted in human subjects, in which the clinical and statistical description, presentations, and analyses are fully integrated into a single report"[4]
- ICH (International Conference on Harmonisation) E3: ICH guidelines on the structure and content of clinical study reports[3]
- Adverse event (AE): "Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have a causal relationship with this treatment"[4]
- Serious adverse event (SAE): "Any untoward medical occurrence that at any dose: results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect"[4]
- Summary tables: in a CSR "All adverse events occurring after the initiation of study treatment . . . should be displayed in summary tables . . . The tables should list each adverse event, the number of patients in each treatment groups in whom the event occurred, and the rate of occurrence"[3]
- Narratives: in a CSR "There should be brief narratives describing each death, each other serious adverse event, and those of the other significant adverse events that are judged to be of special interest because of clinical importance. These narratives can be placed either in the text of the report or in section 14.3.3, depending on their number. Events that were clearly unrelated to the test drug/investigational product may be omitted or described very briefly. In general, the narrative should describe the following: the nature and intensity of event, the clinical course leading up to event, with an indication of timing relevant to test drug/investigational product administration; relevant laboratory measurements, whether the drug was stopped, and when; countermeasures; post mortem findings; investigator's opinion on causality, and sponsor's opinion on causality, if appropriate."[3] Narratives are based on extracted data from source files (for example, case report forms). They are written by medical writers. Narratives can be written before data are finalised, but updates are required based on the final data[5]
- Appendices: CSRs include appendices on study information (for example, protocol and protocol amendments, sample case report forms, a list of institutional review boards/ethics committees, a list of investigators) and patient data listings (discontinued patients, protocol deviations, patients excluded from the efficacy analysis, individual efficacy response data, adverse event listings, individual laboratory measurements listings). Under Directive 2001/83/EC and ICH E3, these appendices do not necessarily have to be submitted to the EMA as part of the regulatory submission for marketing authorisation, but the sponsor must make these available to the EMA on request. The "Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications" lists the appendices required to be submitted to the EMA with each CSR. These appendices include the protocol and amendments to the protocol[3 6 7]
- Individual patient adverse event listings: All adverse events for each patient, including the same event on several occasions, should be available as an appendix of the CSR. ICH E3 suggests the variables, such as patient identifier, the adverse event (preferred term and reported term), duration of the adverse event, severity (for example, mild, moderate, severe), seriousness (serious/non-serious), action taken (none, dose reduced, treatment stopped, etc), and outcome, that should be included in the listing[3]

website http://purl.bioontology.org/ontology/CST; MedDRA versions 2.1 to 16.0 were accessed electronically through an academic subscription.

For all trials we attempted to reconcile each suicidality event in the three data formats. Firstly, using the patient's trial identification number we were able to reconcile data reported in the patient listings with those in the narrative. Secondly, using data (treatment assignment, coded term, and timing of event) from the patient listings and narratives, we were able to reconcile data from these two formats with the data in summary tables.

Results

Six trials (1586 patients) used the coding dictionary COSTART (version number not provided) and three trials (1292 patients) used MedDRA (version 5.0 or 6.0). Adverse events listings for individual patients were available for all nine trials (1672 patients receiving duloxetine, 777 receiving placebo, 70 receiving fluoxetine, and 359 receiving paroxetine). These listings provided data on individual adverse events experienced by each patient in the trial and included the verbatim term, severity of the event, if the adverse event was serious or led to discontinuation of the study drug, and whether the adverse event was considered to be related to the study drug (see table 3). The listings did not, however, provide the preferred term of the event. Narratives were the only data format to provide both verbatim and preferred terms. We were therefore only able to compare verbatim terms to coded terms for those patients who had a narrative—that is, patients who experienced a serious adverse event, discontinued the study drug as a result of an adverse event, or had a clinically significant non-serious adverse event. A median of 11% of patients in each trial had a narrative. We also noted that the listings contained no information on action taken with the study drug in response to the adverse event—for example, dose reduction, the date that the study drug was stopped, whether the adverse event resolved on reducing the dose or stopping the drug, or whether the patient received any treatment for the adverse event.

Within the clinical study report, summary tables of adverse events for some of the trial phases were presented; the lead-in phase of 3-10 days without drugs was always missing. If patients experienced a specific adverse event in the randomised phase, its incidence for each arm was reported in the table. All events were presented as the preferred term.

Individual patient listings versus narratives

The listings of adverse events for individual patients described three suicides and three definitive suicide attempts, which were serious adverse events. The listings also showed a "possible suicide attempt," which was mild, non-serious, and did not result in the patient discontinuing the study drug.

There were narratives for the suicides and definitive suicide attempts, as these were all serious adverse events. From the narratives it could be discerned that verbatim terms were coded to identical terms.

No narrative was present for the patient who experienced a "possible suicide attempt," because the patient did not experience any serious or clinically important adverse events and did not discontinue the study drug as a result of an adverse event. Furthermore, there were no events, such as "overdose" mentioned for this patient in the individual patient listings that could possibly constitute a suicide attempt. Therefore, we did not have any information as to what the possible suicide attempt comprised.

The patient listings described 10 patients who experienced events relating to suicidal ideation (six receiving duloxetine, two receiving placebo, two receiving paroxetine), one patient receiving placebo who experienced "increased suicidality," and one patient receiving duloxetine who experienced "suicide threat" (see supplementary table 1).

Narratives were only available for six patients from three trials (as the other six patients did not experience any serious or clinically important adverse events or adverse events that led to discontinuation). In two of the three trials, COSTART was used. Narratives from these two trials showed that one event of "suicidal urges" while receiving duloxetine and two events of "suicidal ideation" while receiving paroxetine were coded as depression. According to the definitions of the International Conference on Harmonisation, adverse events can include pre-existing conditions that worsen after starting the study drug.[23] One of the two patients receiving paroxetine had a mild baseline "suicidal ideation" that worsened in severity in the randomised phase of the trial (see supplementary table 1), therefore meeting the criterion of an adverse event. However, "suicidal ideation" was only recorded in the narrative as a pre-existing condition, not as an adverse event. Furthermore, there was no mention in the narrative text of a worsening of the severity of suicidal ideation. The only suicidality preferred term in the last version of COSTART (version 5) is suicide attempt. There is no exact lowest level term for suicidal ideation in COSTART; the closest possible matching term is suicidal tendency, which is coded to the preferred term depression.

In the third trial, MedDRA version 6.0 was used, and two events of suicidal ideation were coded as the preferred term suicidal ideation. The event "suicidal threat" (the patient threatened to harm herself while in possession of a knife) was coded as suicidal ideation. Although more recent versions of MedDRA have an appropriate term (lowest level term preparatory actions towards imminent suicidal behaviour, which codes to the preferred term suicidal behaviour), version 6.0 did not.

We also found a case of suicidal ideation (see supplementary table 1) that did not appear in the patient listings but in the narrative text only, in a patient receiving duloxetine who experienced "worsened depression." This finding agrees with the common coding convention of only coding a definitive diagnosis and not its symptoms.

Individual patient listings and narratives versus tables

The three suicides could clearly be identified in the coded data presented in summary tables of adverse events in the clinical study reports. The three definitive suicide attempts and one "possible suicide attempt" came from one trial, and its summary table reported four suicide attempts.

Summary tables showed important loss of information on adverse events. Two of the 10 events related to suicidal ideation were coded as suicidal ideation using MedDRA version 6.0. We were only able to reconcile verbatim terms to coded terms for three of the nine other events in summary tables of trials using the COSTART dictionary.

In all three cases the original term reported by the investigator was coded to the COSTART preferred term depression. Two of these cases (one patient receiving duloxetine and the other receiving paroxetine) occurred in the randomised phase of one trial, and the summary table reported them as depression. The third case (patient receiving paroxetine) occurred in the randomised phase of a different trial where the summary table for the randomised phase reported depression while receiving paroxetine.

The event of "suicidal threat," where a patient receiving duloxetine threatened to harm herself while in possession of a knife, was coded to the preferred term suicidal ideation using MedDRA version 6.0, which was also the term used in the summary table.

We also found instances where events of suicidal ideation were present in patient listings but were absent from summary tables, and vice versa. In the patient listings of one trial there was an adverse event of suicidal ideation in a patient receiving paroxetine that met the criteria of a treatment emergent adverse event in the patient listings, but was only shown as a pre-existing condition, coded to depression, in the narratives. In the summary table there were zero adverse events of depression in the paroxetine arm. Furthermore, in one trial, which used the coding dictionary MedDRA, summary tables of coded data for the open label single arm run-in phase reported three events of suicidal ideation. From the patient listings and narratives, however, we could only identify two adverse events of suicidal ideation.

Discussion

We wanted to assess the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports. From the small number of suicidal events that we were able to reconcile, coding was both accurate, given the constraints of the dictionaries used, and consistent. The suicides were clearly identifiable in all formats of adverse events data whereas, in line with common coding conventions, suicide attempts in tables included both definitive and provisional diagnoses. However, some events of suicidal ideation and preparatory behaviour were obscured in tables owing to the lack of specificity in the coding dictionary used. Instances of suicidal ideation events were present in patient listings but were absent from summary tables, and vice versa. One event of suicidal ideation appeared in the narrative text only. This may result from the common coding convention that if symptoms and a definitive diagnosis are both provided, only the diagnosis is coded.

Strengths and limitations of this study

Our study is based on a small number of trials for a single drug manufactured by a single company.

Another limitation is that, although the guideline for clinical study reports suggests that listings of adverse events for individual patients should include coded terms in addition to the verbatim terms,[3] this was not the case for the nine trials we examined. Our analysis of discrepancies in adverse events data was therefore limited to comparing data already coded in tables to those of narratives, which included verbatim and coded terms, of those patients who had adverse events that were serious, led to discontinuation of the study drug, or were non-serious but clinically important.

Comparisons with other studies

Problems with terms in COSTART, including lack of specific preferred terms, were acknowledged in journal articles in the 1990s (the last version of COSTART was released in 1995).[24 25] It is therefore possible that our finding that adverse events of suicidal ideation were obscured in summary tables of COSTART coded data could also apply to other types of adverse events.

Conclusion and implications for researchers and clinicians

Our study has shown that researchers and clinicians need to be aware that because of coding dictionaries and coding conventions used, adverse events data presented in summary tables may obscure adverse events of importance. Furthermore, important data, in particular the verbatim terms of adverse events, can be presented in the patient listings in the appendices

of clinical study reports. To obtain a more accurate estimate of the incidence of specific adverse events, the verbatim terms should be recoded with the latest version of MedDRA. This is in agreement with informal advice from the FDA and from the MedDRA Maintenance and Support Services Organization. However, researchers contemplating using clinical study reports as sources of data for adverse events need to be aware that access to verbatim terms may not be possible, as the individual patient listings of all adverse events, in contrast with serious adverse events, is not a mandatory part of the submission to the EMA.[6] Furthermore, patient listings may not contain information on certain events (owing to coding conventions), or important information, such as action taken with the study drug and treatments given in response to adverse events. Important adverse events data may therefore only be available in the narratives of patients who experienced adverse events that were serious, led to discontinuation of the study drug, or were non-serious but clinically important.

It should also be noted that, while clinical study reports contain detailed data on adverse events, there is evidence from FDA analyses and court cases that access to case reports forms reveal discrepancies that would not be apparent from clinical study reports alone.[26 27] For example, an FDA analysis of a sample of case report forms from the RECORD trial revealed many missing cases of cardiac problems, which allowed the determination that, in contrast to the manufacturer's (GlaxoSmithKline) claims, rosiglitazone increased the risk of cardiac problems fourfold.[26] Furthermore, case report forms are sometimes unavailable to, or rarely used by, academic authors of journal articles reporting industry sponsored trials. Readers of journal articles should therefore be aware that academic authors often only use data files of coded data or coded data from clinical study reports to perform or check analyses presented in journal articles.[28] Case report forms are currently unavailable to independent researchers. Should case report forms become available, any independent research using case report forms is likely to be costly, in terms of both time and money, as a case report form for a single patient can be hundreds of pages in length and require a considerable infrastructure to ensure unbiased judgments.[26]

In conclusion, adverse event data in tables in clinical study reports may not accurately represent the underlying patient data owing to medical coding dictionaries and coding conventions used. In clinical study reports, the individual patient listings of harms and narratives of adverse events can provide important additional data, including the original terms for adverse events reported by the investigators, which can enable a more accurate estimate of harms.

We thank Julie Borring, Kristine Rasmussen, and Trine Gro Saida for assistance with initial data extraction; the EMA for providing the material and for responding to queries related to the material; and Jesper Krogh for sharing material he obtained from the EMA.

Contributors: All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. EM and BT contributed to the study concept and design. EM, BT, AH and AL contributed to the acquisition of data. All authors contributed to the analysis and interpretation of data, and drafts of manuscripts. All the authors critically reviewed the manuscript for publication. PCG provided administrative, technical, and material support, and was the study supervisor and guarantor.

Funding: This study is part of a PhD (EM) funded by Rigshospitalets Forskningsudvalg. The funding source had no role in the design and conduct of the study; data collection, management, analysis, and interpretation; preparation, review, and approval of the manuscript; or the decision to submit the paper for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: this study is part of a PhD funded by Rigshospitalets Forskningsudvalg; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: The clinical study reports we used can be obtained from the authors (em@cochrane.dk).

Transparency: The manuscript's guarantor affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

1 Doshi P, Jefferson T, Del Mar C. The imperative to share clinical study reports: recommendations from the Tamiflu experience. *PLoS Med* 2012;9:e1001201.
2 Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev* 2012;1:CD008965.
3 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Structure and content of clinical study reports: E3. 1995. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.
4 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Guideline for good clinical practice: E6 (R1). 2013. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1__Guideline.pdf.
5 Challenges of writing patient/ safety narratives—part 1. 2014. http://info.quanticate.com/bid/56497/Challenges-of-Writing-Patient-Safety-Narratives-Part-1.
6 Committee for medicinal products for human use (CHMP). Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications (CHMP/EWP/2998/03 Final). 2004. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003638.pdf.
7 Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the community code relating to medicinal products for human use. 2013. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:311:0067:0067:EN:PDF.
8 Chow S-H, Lui J-P. Design and analysis of clinical trials: concepts and methodologies. 2nd ed. Wiley-Interscience, 2004.
9 Brown EG. Dictionaries and coding in pharmacovigilance. In: Talbot J, Waller P, eds. Stephens detection of new adverse drug reaction. 5th ed. Wiley, 2004:533-57.
10 White CA. A preliminary assessment of the impact of MEDDRA on adverse event reports and product labelling. *Drug Inf J* 1998;32:347-62.
11 Kornai A, Stone L. Automatic translation to controlled medical vocabularies. In: Abraham A, Jain LC, van der Zwaag BJ, eds. Innovation in intelligent systems. Springer, 2004:413-34.
12 MedDRA. History. 2013. www.meddra.org/about-meddra/history.
13 Brown DR, Brown EG, Moulvad TB. A comparison of two medical terminologies in coding and analysing clinical trial safety data. *Int J Pharm Med* 1997;11:85-9.
14 ICH Secretariat. MedDRA term selection: points to consider release 4.6. 2014. www.meddra.org/sites/default/files/guidance/file/9491-1610_termseltc_r4.6_sep2013.pdf.
15 Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. *Am J Psychiatry* 2007;164:1035-43.
16 Peters PH, Gravenstein S, Norwood P, De Bock V, Van Couter A, Gibbens M, et al. Long-term use of oseltamivir for the prophylaxis of influenza in a vaccinated frail older population. *J Am Geriatr Soc* 2001;49:1025-31. [Correction in *J Am Geriatr Soc* 61:478.]
17 Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 2012;344:e1119.
18 Scientific Discussion EMA 2005. 2013. www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000572/WC500036776.pdf.
19 Central authorisation of medicines. 2014. www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000109.jsp.
20 Selective serotonin reuptake inhibitors and serotonin and noradrenaline reuptake inhibitors information for healthcare professionals. 2014. www.mhra.gov.uk/Safetyinformation/Generalsafetyinformationandadvice/Product-specificinformationandadvice/index.htm.
21 ABBYY FineReader 11 Professional edition—full feature list. 2013. http://finereader.abbyy.com/professional/full_feature_list/.
22 Hammad TA, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006;63:332-9.
23 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Statistical principles for clinical trials: E9. 1998. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf.
24 Fizames C. How to improve the medical quality of the coding reports based in Who-Art and Costart use. *Drug Inf J* 1997;31:85-92.
25 Matsumoto C. Implementation and use of the Costart dictionary at Lilly: would we do it again. *Drug Inf J* 1991;25:201-3.
26 US FDA. Advisory Committee Meeting for NDA 21071 Avandia (rosiglitazone maleate tablet). 2010.
27 Healy D. Let them eat Prozac. New York University Press, 2004.
28 Lundh A, Krogsbøll LT, Gøtzsche PC. Access to data in industry-sponsored trials. *Lancet* , 2011;378:1995-6.

What is already known on this topic

For statisticians to analyse adverse events recorded in a clinical trial, it is necessary that events described by the original investigators are coded to terms in a specialised medical coding dictionary

Miscoding of harms can prevent an accurate risk assessment of harms

Extensive coded data on adverse events, in different summaries and tabulations, and provision of original investigator reported terms, can be found in clinical study reports submitted in drug licensing applications to the regulatory authorities

What this study adds

The use of coding dictionaries and coding conventions may inadvertently obscure events that are important in summary tables

Individual patient listings of harms and narratives of adverse events can provide important additional data, including original investigator reported descriptions of the adverse events, which can enable a more accurate estimate of harms

Accepted: 5 May 2014

Cite this as: *BMJ* 2014;348:g3555

Tables

Table 1| MedDRA hierarchical structure

No of terms	Level of term	Example
26	System organ class (SOC)	Psychiatric disorder
>330	High level group terms (HLGT)	Suicidal and self injurious behaviour NEC
>1700	High level terms (HLT)	Suicidal and self injurious behaviour
>20 000	Preferred terms (PT)	Suicidal ideation
>70 000	Lowest level terms (LLT)	Suicidal tendency, active suicidal ideation, death wishes, life weariness, passive suicidal ideation, suicidal ideation, suicidal intention, suicidal plans

MedDRA=Medical Dictionary for Regulatory Activities; NEC=not elsewhere classified.

Table 2| COSTART hierarchical structure[8 9]

No of terms	Level of term	Example
12	Body system	Nervous system
	Mid-level classification	Depression
About 1200	Coding symbol (analogous to preferred term)	Depression
>6000	Glossary terms (used to assist in selection of coding symbol, analogous to lowest level terms of MedDRA)	Suicidal tendency, depression agitated, depression mental, melancholia, depression worsened, depressed reaction, oppression, dejection emotional, moroseness, depression reactive, dysphoria, depression aggravated, depression functional, depression psychic, depression neurotic, anhedonia, depressed state

COSTART=Coding Symbols for a Thesaurus of Adverse Reaction Terms.
Examples using http://bioportal.bioontology.org/ontologies/COSTART.

Table 3| Variables in individual patient listings for adverse events, and in narratives

Variables	ICH E3 specified variables for IP AE listing*	Variables in DLX CSR IP AE listing	ICH E3 specified variables for narrative*	Variables in DLX CSR narratives
Investigator	Yes	Yes	—	—
Treatment group	Yes	Yes		Yes
Patient identifier	Yes	Yes	Yes	Yes
Age, race, sex, weight	Yes	—	Yes	Yes
Location of CRFs, if provided	Yes	—	—	—
The adverse event (preferred term, reported term)	Yes	Reported term only	—	Yes (preferred and reported terms)
Duration of adverse event	Yes	Yes*	—	Yes
Severity	Yes	Yes	—	Only for adverse events that led to discontinuation
Seriousness (serious/non-serious)	Yes	Yes	—	Yes
Action taken (none, dose reduced, treatment stopped, specific treatment instituted, etc)	Yes	—	—	Yes†
Outcome (for example, CIOMS format)	Yes	Start and stop dates of event provided	—	Yes*
Causality assessment (for example, related/not related)	Yes	Yes	Yes	Yes
Date of onset or date of clinic visit at which event was discovered	Yes	Yes	—	Yes
Timing of onset of adverse event in relation to last dose of test drug/investigational product (when applicable)	Yes	Yes	—	Yes
Study treatment at time of event or most recent study treatment taken	Yes	—	—	Yes
Test drug/investigational product dose in absolute amount, mg/kg or mg/m² at time of event	Yes	—	—	Yes
Drug concentration (if known)	Yes	—	—	—
Duration of test drug/investigational product treatment	Yes	—	—	Yes
Concomitant treatment during study	Yes	—	—	Yes
Date study drug started	—	Yes	Yes	Yes
Relevant concomitant/previous illnesses with details of occurrence/duration	—	Yes	Yes	Yes
Relevant concomitant/previous drug with details of dosage	—	—	Yes	Yes
Relevant laboratory measurements	—	—	Yes	Yes

CIOMS=Council for International Organizations of Medical Sciences guidelines; CRF=case report form; CSR=clinical study report; DLX=duloxetine; ICH E3=International Conference on Harmonization of technical requirements for registration of pharmaceuticals for human use: structure and content of clincal study reports: E3; IP AE=individual patient adverse event.

*Can be calculated from information in data format.

†Apparent from information in text.