

vigiGrade: A Tool to Identify Well-Documented Individual Case Reports and Highlight Systematic Data Quality Issues

Tomas Bergvall · G. Niklas Norén ·
Marie Lindquist

Published online: 17 December 2013
© Springer International Publishing Switzerland 2013

Abstract

Background Individual case safety reports of suspected harm from medicines are fundamental to post-marketing surveillance. Their value is directly proportional to the amount of clinically relevant information they include. To improve the quality of the data, communication between stakeholders is essential and can be facilitated by a simple score and visualisation of the results.

Objective The objective of this study was to propose a measure of completeness and identify predictors of well-documented reports, globally.

Methods The Uppsala Monitoring Centre has developed the vigiGrade completeness score to measure the amount of clinically relevant information in structured format, without reflecting whether the information establishes causality between the drug and adverse event. The vigiGrade completeness score (C) starts at 1 for reports with information on time-to-onset, age, sex, indication, outcome, report type, dose, country, primary reporter and comments. For each missing dimension, a penalty is detracted which varies with clinical relevance. We classified reports with $C > 0.8$ as

well-documented and identified all such reports in the WHO global individual case safety report database, VigiBase, from 2007 to January 2012. We utilised odds ratios with statistical shrinkage to identify subgroups with unexpectedly high proportions of well-documented reports. **Results** Altogether, 430,000 (13 %) of the studied reports achieved $C > 0.8$ in VigiBase. For VigiBase as a whole, the median completeness was 0.41 with an interquartile range of 0.26–0.63. Two out of three well-documented reports come from Europe, and two out of three from physicians. Among the countries with more than 1,000 reports in total, the highest rate of well-documented reports is 65 % in Italy. Tunisia, Spain, Portugal, Croatia and Denmark each have rates above 50 %, and another 20 countries have rates above 30 %. On the whole, 24 % of the reports from physicians are well-documented compared with only 4 % for consumers/non-health professionals. Notably, Denmark and Norway have more than 50 % well-documented reports from consumers/non-health professionals and higher rates than for physicians. The rate of well-documented reports for the E2B format is 11 % compared with 22 % for the older INTDIS (International Drug Information System) format. However, for E2B reports entered via the WHO programme's e-reporting system VigiFlow, the rate is 29 %.

Conclusion Overall, only one report in eight provides the desired level of information, but much higher proportions are observed for individual countries. Physicians and e-reporting tools also generate greater proportions of well-documented reports overall. Reports from consumers/non-health professionals in specific regions have excellent quality, which illustrates their potential for the future. vigiGrade has already provided valuable information by highlighting data quality issues both in Italy and the USA.

Electronic supplementary material The online version of this article (doi:10.1007/s40264-013-0131-x) contains supplementary material, which is available to authorized users.

T. Bergvall (✉) · G. N. Norén · M. Lindquist
Uppsala Monitoring Centre, WHO Collaborating Centre
for International Drug Monitoring, Uppsala, Sweden
e-mail: tomas.bergvall@who-umc.org

G. N. Norén
Department of Mathematics, Stockholm University,
Stockholm, Sweden

1 Background

Identification, management and prevention of adverse drug reactions (ADRs) and other patient harms related to the use of medicines rely on sound decision making based on the best available evidence [1]. Since knowledge about the real-life benefits and harms of medicines is limited at launch, the evidence base has to be gradually built up by gathering and analysing relevant clinical patient data throughout the life-cycle of a product. To avoid unnecessary suffering, signals of potential new medication problems need to be identified and communicated as early as possible. A key component of a pharmacovigilance system is its ability to capture and process real-life patient data in a way that allows intelligent and critical analysis and interpretation. To improve the quality of the data, communication between the stakeholders is essential and can be facilitated by a simple score and visualisation of the results.

For many years, individual case reports of suspected harm from medicines collected in post-marketing pharmacovigilance reporting systems have been the basis for signal detection and early risk assessment of potential new medication problems [2, 3]. Their value for signal detection and evaluation is directly proportional to the amount of clinically relevant information they include [4, 5].

Data sources such as electronic healthcare records, claims databases and registries today contribute a useful resource for hypothesis testing and, in some cases, also signal detection [6]. However, individual case report data are still key for hypothesis generation in pharmacovigilance.

1.1 WHO Programme and VigiBase

The WHO Programme for International Drug Monitoring was set up in 1968 with the aim to ensure that early signs of previously unknown medicine-related safety problems are seen. The safety problems would be identified from pooled data, and information about them shared and acted upon by the national pharmacovigilance centres participating in the programme.

The WHO programme's signal detection process is based on data stored in the WHO global individual case safety report database, VigiBase, managed by the WHO Collaborating Centre for International Drug Monitoring, Uppsala Monitoring Centre (UMC). With more than 100 countries contributing data on a regular basis, VigiBase has a unique global coverage enabling international signal detection and assessment, as well as analysis of inter-country or inter-regional reporting patterns.

Following a major database overhaul starting in the late 1990s, VigiBase has been fully compatible with the International Conference on Harmonisation (ICH) E2B [7] standard format since 2002.

1.2 The Need for Data Quality Management

Major regulatory agencies have pointed to the need for quality management systems as an essential component of good pharmacovigilance practices [8, 9]. The guidance documents address the quality of the pharmacovigilance processes: data collection, storage, management, risk assessment and communication. Whilst quality processes are essential to ensure the integrity and validity of the information through the data processing cycle, the reliability and credibility of assessments based on reported information also depend on the quality of the data itself. Therefore, a robust pharmacovigilance quality management system should include efforts to assess and improve data quality [10].

In its guidance on pharmacovigilance quality systems for European authorities, the pharmaceutical industry and the EudraVigilance database, the European Medicines Agency (EMA) requests that procedures and processes be established and maintained in order to ensure the evaluation of the quality, including completeness, of pharmacovigilance data submitted [9].

In *Best Practice in Reporting of Individual Case Safety Reports (ICSRs)* [11], specific examples of good practice are highlighted, with the aim of contributing to consistent and high-quality data.

Ideally, in a quality management system all of the quality parameters given in Table 1 should be considered. To base a safety signal only on poorly documented case reports without the necessary information needed to make a clinical assessment and exclude obvious confounding is, and should be, open to criticism. Problems related to completeness (i.e. missing data) have long been identified as important factors hampering the usefulness of existing individual case report data. A study undertaken in 2000 showed that less than half of the reports in VigiBase contained even basic information such as reaction onset and medicine treatment dates, and only a small fraction (11.5 and 10.6 % in 1995 and 2000, respectively) included dates as well as indication for treatment and patient outcome [12]. The introduction of the much more extensive E2B reporting format, replacing the original WHO reporting format, has not eliminated the problem of missing data. The E2B format describes the structure of an ideal data set, and thus in theory supports the collection of all information needed for a thorough clinical assessment. However, like the old reporting format, it is a way of describing how to structure and format information (addressing conformity)—it does not in itself enforce completeness, nor does it give guidance as to which data items are most important for clinical case assessment.

An early attempt to provide such guidance was made in 1990, when a paper was published that devised criteria for

Table 1 Quality parameters that should be considered in a complete quality management system

Parameter	Description
Accuracy	Does the information represent the ‘real-world’ values correctly?
Completeness	Are all critical items included? Are they recorded in a usable way?
Conformity	Do the data values conform to specified formats and controlled vocabularies?
Consistency	Do interdependent data items provide conflicting information?
Currency	Is the information up to date?
Duplication	Does the data set contain multiple representations of the same case?
Integrity	Can related records be linked together?
Precision	Is the output level of detail supported by input data?
Relevance	Is the data fit for purpose?
Understandability	Can the data be interpreted correctly? No ambiguities?

the amount and type of information needed to produce a well-founded early ADR signal from VigiBase [4]. Based on the results in this paper, a new field ‘documentation grading’ was added to VigiBase. The grading has been used to facilitate the identification of well-documented cases for clinical review, but has also been useful to identify and rectify problems related to missing data in the reports received.

More recently, Agbabiaka et al. [13] proposed a structured assessment of the quality of individual case reports in scientific publications. Their questionnaire is very comprehensive and is intended as a support for manual review but not for automated assessment.

1.3 Aim

The aim of this study was to propose a measure of completeness and identify predictors of well-documented individual case safety reports, globally. In revising the UMC quality management system for VigiBase, we will extend the data quality grading to include a more sophisticated measure of completeness of clinically relevant information in structured format.

2 Methods

We propose the vigiGrade completeness score, which measures the amount of information available in structured format on individual case reports. It does not reflect to what extent the information strengthens the suspicion of a causal association between the medicine and adverse event, but instead the dimensions are weighted by their relative importance for causality assessment and follow the principles listed in Table 2. By dimensions, we denote pieces of information about a suspected ADR, such as the time-to-onset or patient age. While these are intended to be generic and apply for any collection of individual reports, their exact implementation will vary with the source database. In

other words, the data elements considered in the evaluation of whether a specific dimension is available or not will vary across databases.

2.1 Scope of Evaluation

The dimensions included in the current implementation of the vigiGrade completeness score are presented in Table 3. Each dimension has an associated penalty that is imposed in its absence. Imprecise information such as a time-to-onset of –1 to 1 month or an age specified as ‘adult’ is penalised, but less so than completely missing information. The dimensions and their associated penalty factors were determined by three UMC pharmacovigilance experts with medical training, through consensus, to match the relative importance of each dimension to causality assessment. Three levels of importance were distinguished: Essential (information without which reliable causality assessment is impossible); Important (information without which reliable causality assessment is very difficult); and Supportive (information that is valuable but without which causality assessment can still typically be performed). The penalties for missing information are the same across each level of importance.

The vigiGrade completeness score is restricted to information that is important to causality assessment and expected to be present on a majority of reports. As an example, lack of information on the outcome of a dechallenge intervention is not penalised, since dechallenge interventions cannot be expected to be performed under all circumstances. For example, it is not possible to dechallenge a vaccination, and for this reason such dimensions were not included.

The content of each database field is checked and information of the wrong type is treated as missing. As an example, ‘XXX’ in the age unit field will be treated as missing information, and so will ‘#’ in the dose duration field. Similarly, the text ‘years’ as age unit on an E2B report will be treated as missing since the correct format is ‘801’ according to the E2B guidelines [7].

Table 2 Principles for the vigiGrade completeness score

Principle	Idea	Example
Measure the amount of information on reports	The score shall reflect the amount of information provided, but not whether it strengthens the suspicion of a causal relation; the magnitude of the score shall reflect the amount of information, and its maximum value shall be 1	Information on time-to-onset is rewarded even if it makes it unlikely that the adverse event is truly caused by the drug
Consider dimensions not data elements	The score shall focus on dimensions, and not on specific data elements. It shall account for the fact that the same information can be provided through different combinations of data elements	For time-to-onset, the reaction start date is only valuable together with the treatment start date, and redundant if time-to-onset is explicitly reported in a separate field
Weigh by importance	The penalties for absence of information on specific dimensions shall reflect the importance of that dimension for causality assessment	Absence of information on time-to-onset receives greater penalty than absence of information on dose
Allow substantial penalties for individual dimensions	Absence even of single dimensions may significantly decrease the score ^a	Absence of information on either (1) time-to-onset or (2) both patient age and sex each decrease the score by 50 % or more
Focus on machine-readable information	The score shall account only for information that can be identified by a computer, for scalability and secondary use	Information on time-to-onset in a free-text case narrative is not rewarded

^a This is the primary motivation for the multiplicative penalties—if missing information would be penalised by subtraction, it would not be possible to penalise individual missing information items to the same extent

Table 3 Overview of the dimensions accounted for in the vigiGrade completeness score

Dimension	Description	Considerations	Penalty (%)
Time-to-onset	Time from treatment start to the suspected ADR	Imprecise information penalised if there is ambiguity as to whether the drug preceded the adverse event; by 30 % if the uncertainty exceeds 1 month, 10 % otherwise See Fig. 1 for examples	50
Indication	Indication for treatment with the drug	Penalty imposed if information is missing or cannot be mapped to standard terminologies such as ICD or MedDRA	30
Outcome	Outcome of the adverse event in this patient		30
Sex	Patient sex	‘Unknown’ treated as missing	30
Age	Patient’s age at onset of the suspected ADR	Age ‘unknown’ treated as missing 10 % penalty imposed if only age group is specified	30
Dose	Dose of the drug(s)		10
Country	Country of origin	Supportive in causality assessment since medical practice and adverse reaction reporting vary between countries	10
Primary reporter	Occupation of the person who reported the case (e.g. physician, pharmacist)	Supportive in causality assessment since the interpretation of reported information may differ depending on the reporter’s qualifications ‘Unknown’ penalised as missing information, but ‘other’ not penalised	10
Report type	Type of report (e.g. spontaneous report, report from study, other)		10
Comments	Free-text information	Uninformative text snippets excluded	10

A detailed description of the included E2B fields and their corresponding fields in VigiBase is provided as Electronic Supplementary Material ADR adverse drug reaction, ICD International Statistical Classification of Diseases and Related Health Problems, MedDRA Medical Dictionary for Regulatory Activities (a registered trade mark belonging to the International Federation of Pharmaceutical Manufacturers Associations)

2.2 Algorithm

The vigiGrade completeness score ranges from 0.07 to 1. It starts at 1, and for every missing dimension, the corresponding penalty factor in Table 3 is applied. For example,

completeness is reduced by 50 % (multiplied by a factor 0.5) if time-to-onset is not available and by 30 % (multiplied by a factor 0.7) if patient age is not specified. The completeness score for a drug–ADR combination on a report is computed as:

Fig. 1 Four examples of penalties (or lack thereof) when there is imprecise or missing information on time-to-onset. *ADR* adverse drug reaction, *TTO* time-to-onset



$$C = \prod_{i=1}^{10} (1 - P_i) = (1 - P_1) \dots (1 - P_{10}),$$

where P_i denotes the penalty according to Table 3 for dimension i (when information is not missing, the penalty is 0). Thus, the maximum completeness is 1 and the minimum completeness is $1 \times 0.5 \times 0.7^4 \times 0.9^5 = 0.07$.

In order to obtain the overall completeness for a report, completeness is first computed for every reported drug-ADR pair, and then aggregated to an average to yield a score for the corresponding report (restricted to drugs listed as suspected or interacting on the report, i.e. excluding drugs listed as concomitant):

$$C = \sum_{j=1}^m \frac{C^j}{m},$$

where j denotes the current drug-ADR combination and m denotes the total number of drug-ADR combinations for the report. An example of how to use the vigiGrade completeness score in practice can be found in Fig. 2.

2.3 Empirical Evaluation

We utilised vigiGrade to identify well-documented reports in VigiBase. For the purpose of this study, we defined reports with completeness >0.8 as well-documented. This threshold requires all of the important and essential dimensions, i.e. $\geq 30\%$ penalty, to be provided and allows at most two of the supportive dimensions, with 10% penalty, to be missing. For VigiBase as a whole, we determined the average completeness, the total number of

well-documented reports and the variation over time in the proportion of well-documented reports.

In a subsequent analysis, we examined the nature of well-documented reports in VigiBase between 1 January 2007 and 4 January 2012 (the most recent 5 years at the time of the study) and how they differ from other reports in VigiBase in the same time period. Covariates associated with well-documented reports were identified based on odds ratios subjected to statistical shrinkage to protect against spurious associations [14]. The strength of shrinkage was set to 1% ¹ of the total number of well-documented reports in this time period, and as a threshold to identify interesting deviations we required the lower limit of a 99% two-sided credibility interval of the odds ratio to exceed 0.5.

The scope of this analysis included country of origin, primary reporter and report format [the current standard E2B versus the older INTDIS (International Drug Information System) format, with a subgroup analysis for the E2B reports originating from the WHO International Drug Monitoring programme's on-line reporting tool, VigiFlow]. The analysis of primary reporter was broken down by country of origin to study differences between national reporting systems. We restricted our review of results to countries with at least 1,000 reports in total in VigiBase within the given timeframe. To enable review and evaluation of the selected weighting scheme, we inspected the completeness of individual dimensions for the three countries with highest completeness, for the different report formats, for physicians and for consumers/non-health professionals, respectively.

¹ This choice of shrinkage provides adaptability to the size of the case series to be analysed and has proven useful in empirical studies; results not shown here.

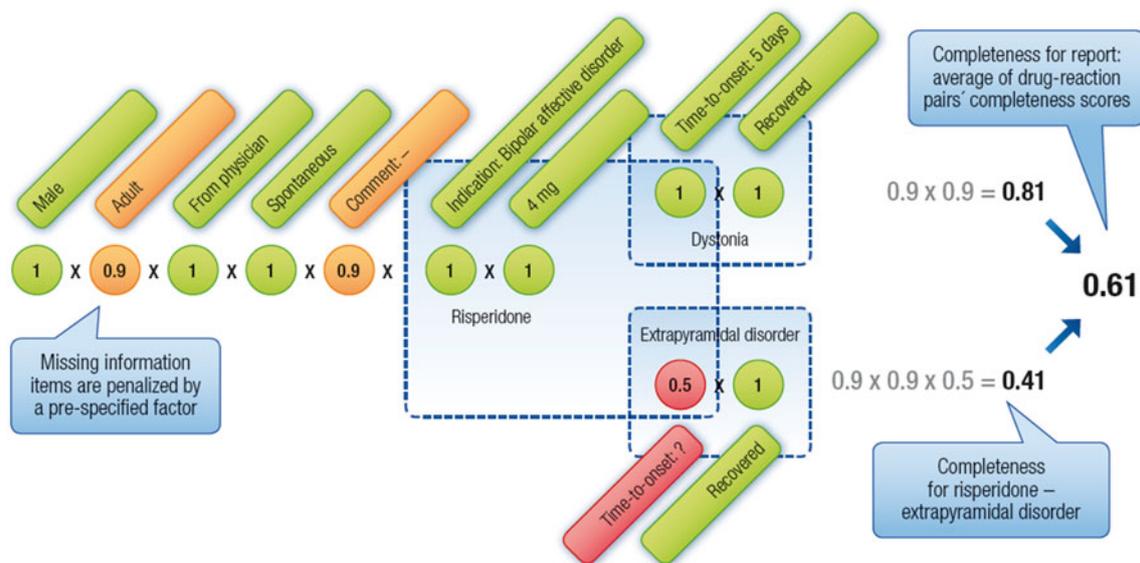


Fig. 2 An example of how the vigiGrade completeness score is calculated for a report

2.4 Prospective Evaluation

During the development of vigiGrade, VigiBase was continuously monitored for quality issues. The actual implementation of vigiGrade varied over time, reflecting the iterative and incremental improvement process. The aim was to discover internal administration problems as well as transmission errors.

3 Results

3.1 Well-Documented Reports in VigiBase Overall

There are a total of 7.0 million reports in VigiBase up until January 2012, having an average completeness of 0.45. 900,000² (13 %) reports have vigiGrade completeness higher than 0.8 and are classified as well-documented for the purpose of our study. Figure 3 shows the variation over time in the proportion of well-documented reports and average completeness in VigiBase (non-cumulative). Both the average completeness and the proportion of well-documented reports were higher until 1980 and have since declined: the average completeness from around 0.50 to 0.45, and the proportion of well-documented reports from around 25 to 13 %.

3.2 Well-Documented Reports in VigiBase Since 2007

Between January 2007 and January 2012, 3.3 million reports were entered into VigiBase. The average

² All numbers in this section are rounded to two significant figures.

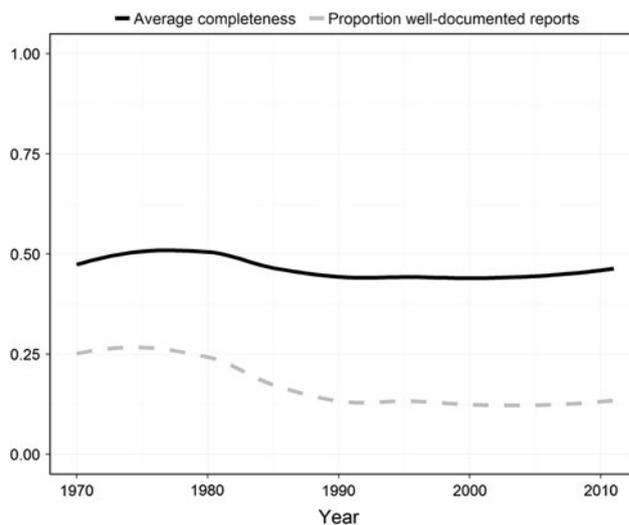
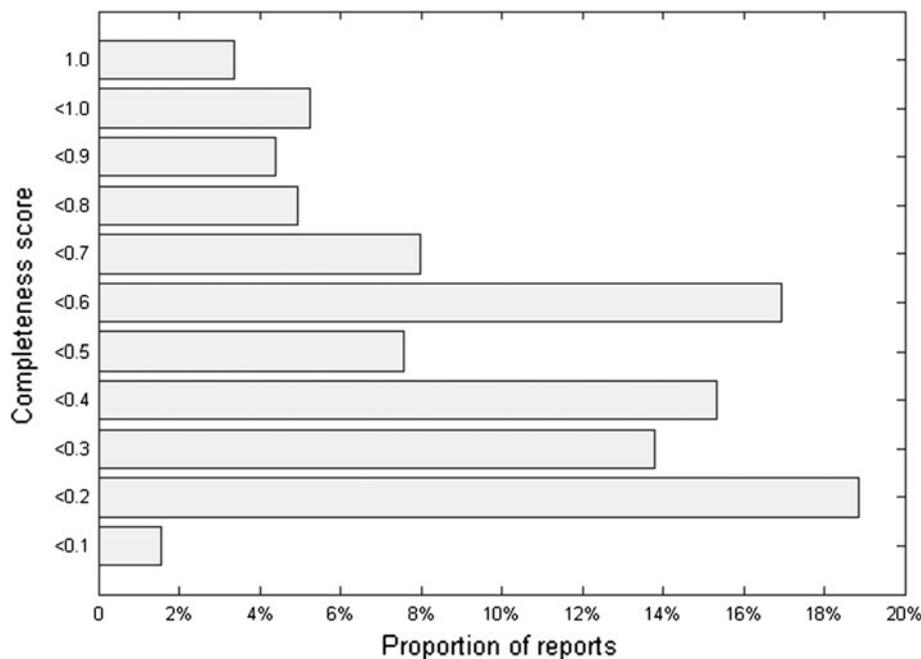


Fig. 3 Distribution of completeness and the proportion of well-documented reports over time in VigiBase

completeness was 0.46 and 430,000 reports (13 %) were classified as well-documented. The median completeness was 0.41 with an interquartile range of 0.26–0.63. Figure 4 shows the distribution. Figure 5 shows the number of well-documented reports per country, in descending order, for countries with at least 1,000 reports in total since 2007. The graph also indicates the expected number of well-documented reports for each country, which is 13 % of the total number of reports for the country. The five countries with the greatest numbers of well-documented reports are Italy, Germany, Spain, Thailand and the USA. Thirty countries had significantly higher than expected numbers of well-documented reports (shrunk log odds ratio exceeding 0.5,

Fig. 4 Empirical distribution of vigiGrade completeness across the 3.3 million reports in VigiBase between January 2007 and January 2012



as described in the Empirical Evaluation section) and these are listed in red in Fig. 5. The highest proportion of well-documented reports in a country with at least 1,000 reports is Italy, with 65 %. Tunisia, Spain, Portugal, Croatia and Denmark each have more than 50 % well-documented reports, whereas another 20 countries have more than 30 % well-documented reports. Altogether, 66 % of the well-documented reports come from Europe, whereas the overall proportion of reports from Europe in this data subset is 23 %.

Figure 6 shows the number of well-documented reports per primary reporter, and Fig. 7 shows the proportion of well-documented reports by country for different types of primary reporter. 69 % of the well-documented reports in VigiBase come from physicians. On the whole, 24 % of the reports from physicians are well-documented compared with 16 % for pharmacists, 14 % for ‘other health professionals’ and only 4 % for consumers/non-health professionals, overall. The variation between countries is substantial, however.

Of the countries with at least 1,000 reports in total and 100 consumer/non-health professional reports, Denmark and Norway both have more than 60 % well-documented reports from consumers/non-health professionals, and Italy and The Netherlands have more than 40 % well-documented reports, whereas no other country has a rate exceeding 30 %.

More than 50 % of the reports from ‘other health professionals’ in Spain, Norway and Italy are well-documented. Altogether, there are 15 countries with at least 1,000 reports in total and 100 reports from ‘other health professionals’, for which at least 30 % of the ‘other health professional’ reports

are well-documented. From Ireland, 27 % of the ‘other health professional’ reports are well-documented compared with only 9 % of the Irish reports overall.

More than 50 % of the reports from pharmacists in Italy, Portugal and Spain are well-documented, and altogether there are 15 countries with at least 1,000 reports in total and 100 reports from pharmacists, for which at least 30 % of the pharmacist reports are well-documented. From India, 47 % of the pharmacist reports are well-documented compared with 31 % of the Indian reports overall.

For Italy, 74 % of the physician reports are well-documented, and so are more than 50 % of the physician reports from Portugal, Venezuela, Tunisia, Spain, Croatia, Denmark and Norway. Altogether, there are 26 countries with at least 1,000 reports in total and 100 reports from physicians, for which at least 30 % of the physician reports are well-documented. From Nigeria, 29 % of the physician reports are well-documented compared with only 10 % of the Nigerian reports overall.

The primary reporter ‘other’ represents a reported field in the old INTDIS format of ‘Not a doctor or dentist’ and the type of reporter could vary between countries. For Sweden, 71 % of these reports are well-documented. Another six countries with at least 1,000 reports in total and 100 reports from others have more than 30 % well-documented reports in this category. From Peru, 35 % of the ‘other’ reports are well-documented compared with only 2 % of reports from Peru overall.

Reports using the E2B format have an average completeness of 0.44 with 11 % well-documented reports compared with an average completeness of 0.53 with 22 % well-documented reports for the INTDIS format. However,

Fig. 5 Countries with at least 1,000 reports in VigiBase between 2007 and 2012 ordered by the number of well-documented reports. *Red text* indicates the 30 countries that had significantly higher than expected numbers of well-documented reports (shrunk log odds ratio exceeding 0.5)

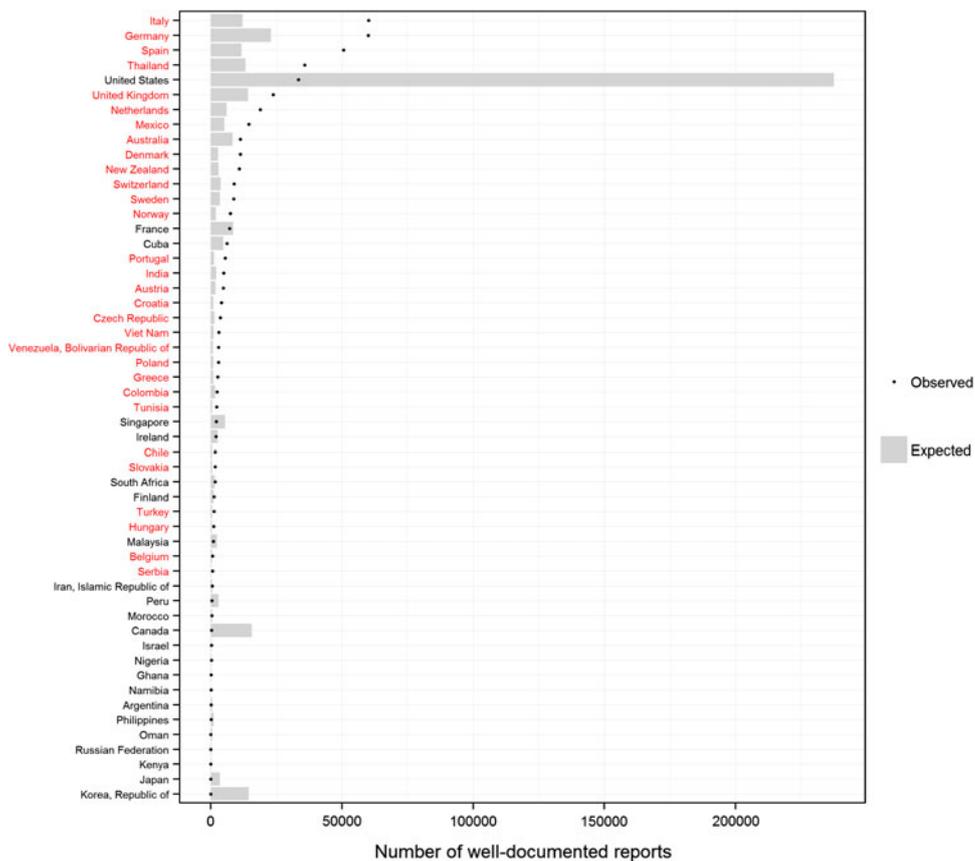
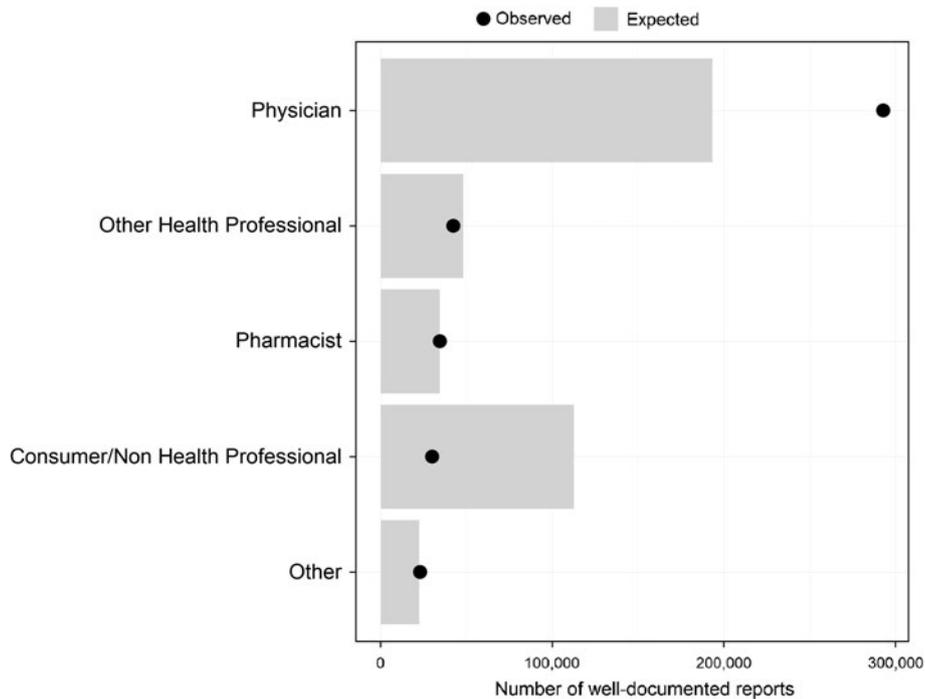


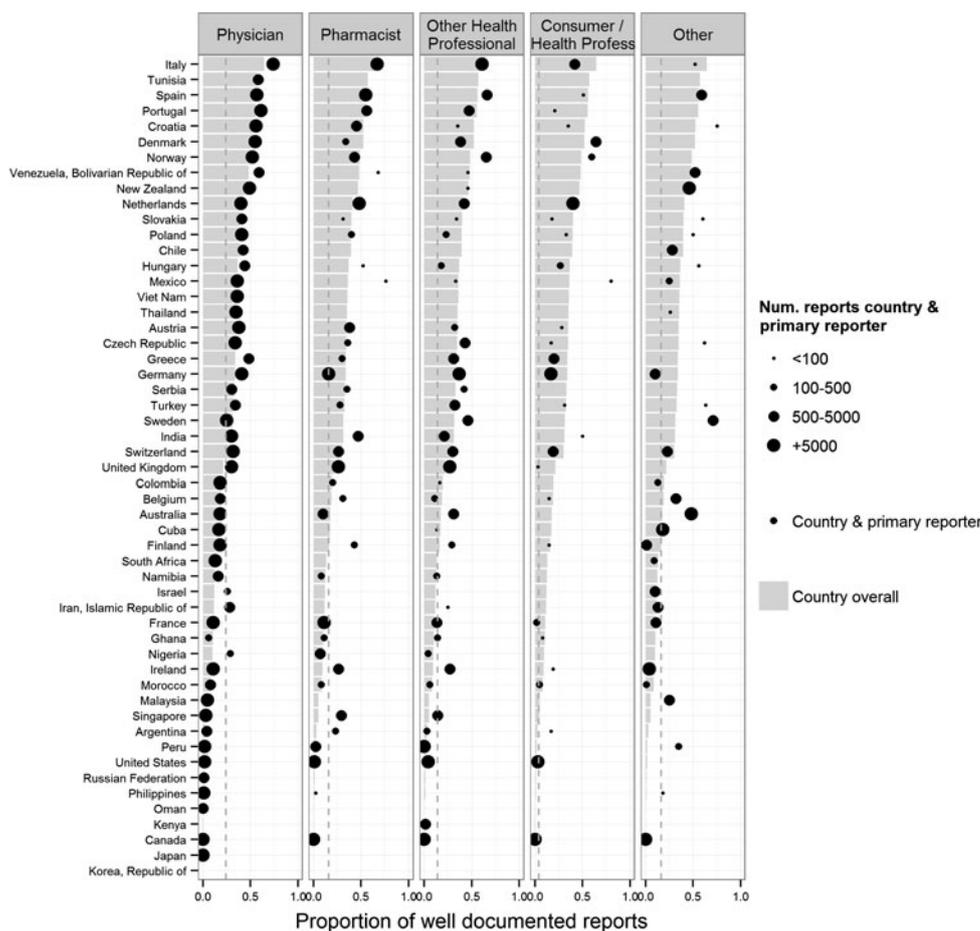
Fig. 6 Number of well-documented reports for different primary reporters



E2B reports via the WHO programme's electronic reporting system VigiFlow have an average completeness of 0.61 and 29 % well-documented reports.

The completeness of individual dimensions for the three countries with highest completeness, for the different report formats, for physicians and for consumers/non-

Fig. 7 The proportion of well-documented reports by country and primary reporter is marked by a *black circle*, the size corresponding to the number of reports. The *grey bars* represent the proportion of well-documented reports for the country overall and the *dotted vertical line* the proportion for the primary reporter overall



health professionals, respectively, is displayed in Table 4. For the three countries of interest, we note that Italy has high completeness for all dimensions, whereas Tunisia and Spain would have suffered from greater penalties on lack of information on dose and free-text comments. Reports on the E2B format carry more information on indication for treatment and free-text comments but less information on patient age, outcome and dose than do reports on the INTDIS format. Reports from VigiFlow carry more information on time-to-onset, dose and free-text comments than other E2B reports and INTDIS reports. Reports from consumers/non-health professionals often lack information on patient age, and carry few free-text comments, but provide information on the indication for treatment slightly more often than reports from physicians.

3.3 Examples of Prospective Discoveries

3.3.1 Miscoded Age Unit from the USA

An unexpected drop in completeness for reports from the US FDA was observed in 2011, as seen in Fig. 8. From

2010 to 2011 the average completeness decreased from 0.45 to 0.30. Subsequent analyses revealed that from 2011 and onwards, the age unit format on reports from the USA did not conform to the E2B guidelines. As a result, all American reports from 2011 to date lacked age information in VigiBase, and none of them were classified as well-documented (since missing age is penalised by 30 %). This issue was communicated to the US FDA and has been addressed in subsequent versions of VigiBase.

3.3.2 Missing Outcome on Italian Reports

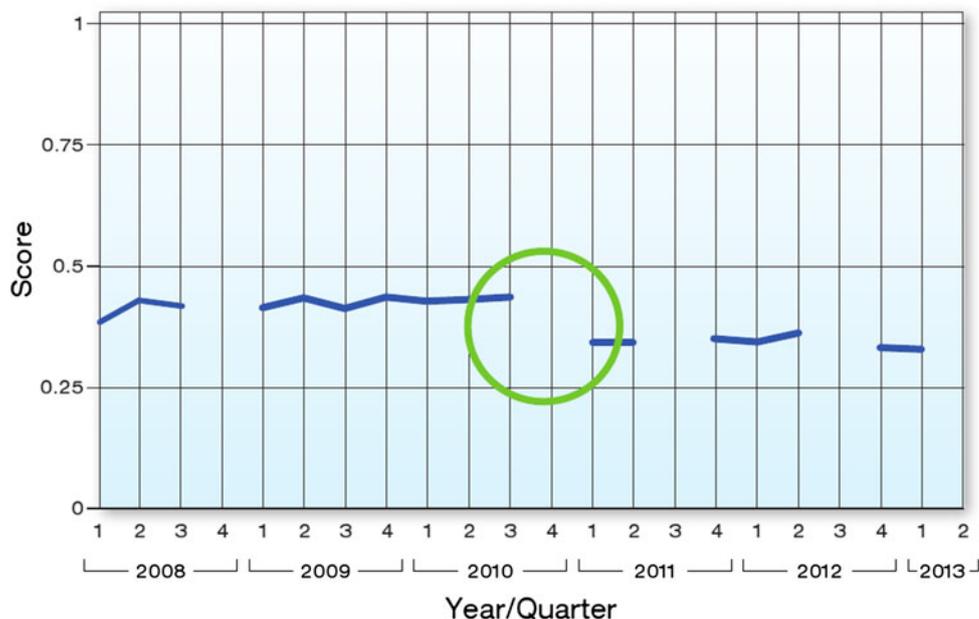
A lower than expected completeness for reports from Italy was observed in 2011. This could be traced to a consistent lack of information on outcome (see Fig. 9). At this time, no Italian reports would have been classified as well-documented. The issue was communicated to the Italian authorities who resubmitted all their reports with the outcome information included. This was done before the initiation of the study at hand. As a result, Italian reports as represented in VigiBase today are the most complete for any country with at least 1,000 reports.

Table 4 Completeness of individual dimensions for selected data subsets

Subset	Time to onset	Age	Indication	Outcome	Primary reporter	Sex	Report type	Country	Dose	Comments
Italy	0.88	0.99	0.82	0.81	1.00	0.99	1.00	1.00	0.82	0.84
Tunisia	0.87	1.00	0.77	0.95	0.99	0.99	1.00	1.00	0.29	0.02
Spain	0.93	0.95	0.75	0.82	0.99	0.98	1.00	1.00	0.50	0.55
INTDIS	0.60	0.92	0.39	0.75	0.92	0.98	1.00	1.00	0.33	0.07
E2B	0.57	0.60	0.52	0.47	0.71	0.97	0.99	1.00	0.22	0.31
VigiFlow	0.71	0.95	0.48	0.75	0.93	0.97	1.00	1.00	0.45	0.80
Consumer/non-health professional	0.46	0.42	0.65	0.49	1.00	0.97	1.00	1.00	0.27	0.03
Physician	0.60	0.74	0.58	0.57	1.00	0.96	0.97	1.00	0.35	0.28

INTDIS International Drug Information System

Fig. 8 The average completeness over time for reports from the USA. A noteworthy decline in 2011 was due to miscoding of the age unit format in their E2B reports. Historical data, not representative of US reports from this time period as represented in VigiBase today



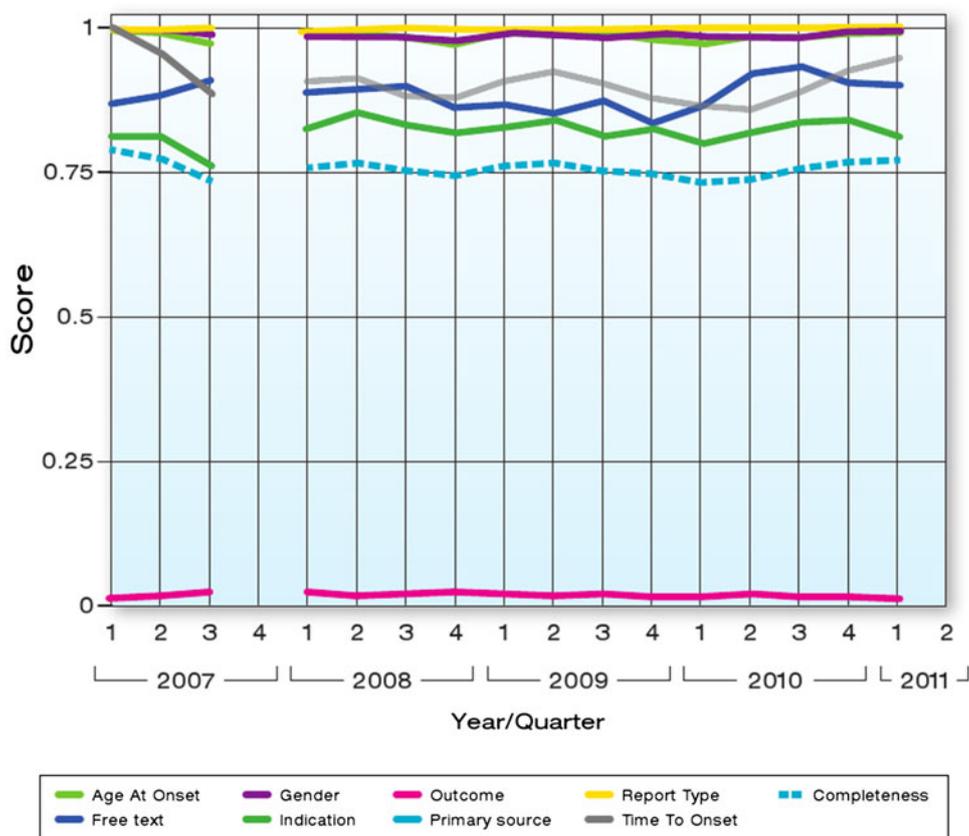
4 Discussion

There are nearly 1 million well-documented reports in VigiBase. These reports all contain the fundamental information required for causality assessment, which includes time-to-onset, patient age and sex, outcome, indication for treatment, and more. Overall, one in eight reports in the past 5 years provide this level of detail, but specific countries perform much better, with rates over 50 %. It is encouraging to see that these same countries maintain high reporting rates per capita, so that a focus on quality does not compromise quantity. With that said, high-quality information comes at a cost in effort and time. The countries with the highest rates of well-documented reports all work hard to obtain the relevant information, and if the initial reports do not suffice, they will contact the health professionals or patients to learn more. In countries such as Italy, Spain and Norway, this work is done at regional

pharmacovigilance centres, whereas in countries such as Croatia it is driven directly by the national centre.

From Denmark, more than three out of five reports from consumers and non-health professionals are well-documented, which is significantly higher than for pharmacists and ‘other health professionals’, and just above the rate for physicians. More strikingly, it far exceeds the overall rate of less than one in 20 well-documented reports from consumers/non-health professionals overall. The Danish Health and Medicines Authority provides a tool for on-line patient reporting, which prioritises ease of use and focuses on ascertaining the key information. This may be a valuable example for others to follow, especially in light of the recent studies showing that direct patient reports can be an excellent complement to reports from physicians and other health professionals [15–17]. In this analysis a distinction could not be made between reports submitted directly by patients and those originating from patients but submitted

Fig. 9 Outcome information as reported by Italy between 2007 and 2011. Historical data, not representative of Italian reports from this time period as represented in VigiBase today



to the national pharmacovigilance centre by a pharmaceutical company. To do this, other E2B fields identifying the sender of the report would have to be used in addition to the primary reporter.

Report quality has declined over time, and it should be investigated whether the regulatory focus on timeliness (15 days rules, etc.) has had an adverse impact on quality. If so, we should carefully consider whether such regulations strengthen or weaken the pharmacovigilance system overall. The decline also seems to coincide with the start of the mandatory reporting by industry and the introduction of a more comprehensive reporting format, which does not in itself improve data quality. As the prospects of signal detection in longitudinal observational databases are improved [18], it becomes even more important to safeguard the unique strengths of individual case reports: their ability to capture just the right information to allow for causality assessment. Initiatives such as the European SALUS project explore the integration of on-line ADR reporting within electronic health record systems, and shall provide insights into the extent with which automatic inclusion of information from health records and requests for the reporter to focus on the clinical assessment may improve quality. Whereas the proportion of well-documented reports is higher for the historical INTDIS format than for the current E2B format, even higher rates are

observed for E2B reports submitted through the WHO programme's electronic reporting tool, VigiFlow. With that said, on-line reporting is neither panacea nor prerequisite for quality—the many well-documented reports in Italy and Spain come out of systems that are largely paper-based.

An earlier version of the vigiGrade completeness score has been in routine use in VigiBase since 2010 (without consideration of dose information and with variations in some of the other penalties) to monitor incoming reports for quality; it was through vigiGrade that systematic processing errors such as the miscoded age units on American reports and the missing outcomes on Italian reports were identified. Measuring and communicating quality is the first step towards better reports, and we hope that feedback to pharmacovigilance professionals will in the end yield better reporting processes and more fit-for-purpose reporting forms. At the other end, vigiGrade should enable safety scientists to home in on the most informative reports in larger series of reports. Presence of information does not guarantee accuracy or relevance, and the information provided may not always strengthen the suspicion of a causal association. Still, it may make sense to start the analysis where there are data to work with. Related to this, we are currently investigating whether well-documented reports may be used as one of the variables in a predictive model for ADR signals.

vigiGrade distinguishes different aspects of quality according to the outline in Lindquist [10]. An important advantage compared with the earlier implementation of documentation grading in VigiBase is that the vigiGrade completeness score considers each dimension in parallel, instead of in sequence: even when information on time-to-onset is lacking, the other dimensions are evaluated and accounted for in the total completeness score. vigiGrade considers many of the same fields as does the structured assessment proposed by Agbabiaka et al. [13], but is less comprehensive. Specifically, it does not evaluate dimensions for which absence of information cannot be distinguished from information on absence in VigiBase. On the other hand, it is a scalable solution that allows automated database-wide analyses. By design, it allows for significant penalties of a variety of missing dimensions.

5 Limitations

vigiGrade measures the amount of information in structured format on reports as represented in VigiBase. Original reports at each national centre may contain more information than is available in VigiBase, and it would be interesting to explore where information is lost along the way. This requires consistent measurement of report quality across multiple data sets, and vigiGrade could provide the basis. Its completeness score can be implemented for any collection of individual case reports, optionally with a different set of dimensions or weights. The weights and ascertainment of dimensions proposed here should be considered as starting points for design parameters that should be further scrutinised and refined, based on feedback from the members of the WHO Programme for International Drug Monitoring. Specifically, one may want to account for the information on concomitant medication in the completeness score or consider additional data elements for some of the dimensions. vigiGrade in its current form focuses primarily on structured data fields, but what truly matters is that the information is computationally accessible. As we develop natural language processing techniques that can extract meaning from free text, the focus of vigiGrade can be expected to shift in the same direction.

6 Conclusions

Overall, only one in eight reports provide the desired level of information, but much higher proportions are observed for individual countries, such as Italy, Tunisia and Spain. Physicians and e-reporting also yield higher proportions of well-documented reports. Reports from consumers and non-health professionals have excellent

quality in specific regions, which illustrates their potential for the future. Future research should explore other aspects of quality, the adaptation of vigiGrade completeness to other data sets, and its use to account for the quality of individual reports in computerised ADR surveillance.

Acknowledgments The authors are indebted to the national centres that contribute data to the WHO Programme for International Drug Monitoring. We would also like to thank Richard Hill, Pia Caduff, Geraldine Hill and Ralph Edwards for expert clinical advice and Jessica Nilsson for expert advice on VigiBase. The opinions and conclusions expressed here are not necessarily those of the national centres, nor of the WHO. No external sources of funding were used to conduct this study or prepare this manuscript. Tomas Bergvall, G. Niklas Norén and Marie Lindquist have no conflicts of interest that are directly relevant to the content of this study.

References

1. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet*. 2000;356(9237):1255–9.
2. van Puijenbroek EP. Case reports and drug safety. *Drug Saf*. 2006;29(8):643–5. doi:10.2165/00002018-200629080-00001.
3. Clarke A, Deeks J, Shakir SW. An assessment of the publicly disseminated evidence of safety used in decisions to withdraw medicinal products from the UK and US markets. *Drug Saf*. 2006;29(2):175–81. doi:10.2165/00002018-200629020-00008.
4. Edwards IR, Lindquist M, Wiholm BE, Napke E. Quality criteria for early signals of possible adverse drug reactions. *Lancet*. 1990;336(8708):156–8.
5. Vandembroucke JP. In defense of case reports and case series. *Ann Intern Med*. 2001;134(4):330–4. doi:10.7326/0003-4819-134-4-200102200-00017.
6. Star K, Bate A, Meyboom RH, Edwards IR. Pneumonia following antipsychotic prescriptions in electronic health records: a patient safety concern? *Br J Gen Pract*. 2010;60(579):385–94.
7. ICH guideline E2B (R2). Electronic transmission of individual case safety reports message specification (ICH ICSR DTD Version 2.1), Final Version 2.3, Document Revision Feb 1, 2001. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM149932.pdf>. Accessed 9 Dec 2013.
8. Guidance for industry: good pharmacovigilance practices and pharmacoepidemiologic assessment [Internet document]. <http://www.fda.gov/downloads/regulatoryinformation/guidances/ucm126834.pdf>. Accessed Jun 2013.
9. Good Pharmacovigilance Practices, Module I—Pharmacovigilance systems and their quality systems [Internet document]. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129132.pdf. Accessed Jun 2013.
10. Lindquist M. Data quality management in pharmacovigilance. *Drug Saf*. 2004;27(12):857–70.
11. Medicines and Healthcare Products Regulatory Agency. Best practice in reporting of individual case safety reports (ICSRs). Internet Document: [55 pages]. Available from: <http://www.mhra.gov.uk/home/groups/comms-ic/documents/publication/con108808.pdf>. Accessed Jun 2013.
12. Lindquist M. Seeing and observing in international pharmacovigilance—achievements and prospects in worldwide drug safety. Nijmegen: University of Nijmegen; 2003.

13. Agbabiaka TB, Savović J, Harris R, Ernst E. The development of a tool to assess the quality of case reports of adverse events. *Int J Risk Saf Med.* 2008;20(3):123–33.
14. Norén GN, Hopstadius J, Bate A. Shrinkage observed-to-expected ratios for robust and transparent large-scale pattern discovery. *Stat Methods Med Res.* 2011;. doi:[10.1177/0962280211403604](https://doi.org/10.1177/0962280211403604).
15. Blenkinsopp A, Wilkie P, Wang M, Routledge PA. Patient reporting of suspected adverse drug reactions: a review of published literature and international experience. *Br J Clin Pharmacol.* 2007;63(2):148–56. doi:[10.1111/j.1365-2125.2006.02746.x](https://doi.org/10.1111/j.1365-2125.2006.02746.x).
16. de Langen J, van Hunsel F, Passier A, de Jong-van den Berg L, van Grootheest K. Adverse drug reaction reporting by patients in the Netherlands: three years of experience. *Drug Saf.* 2008;31(6):515–24.
17. Harmark L, Lie-Kwie M, Berm L, de Gier H, van Grootheest K. Patients' motives for participating in active post-marketing surveillance. *Pharmacoepidemiol Drug Saf.* 2013;22(1):70–6. doi:[10.1002/pds.3327](https://doi.org/10.1002/pds.3327).
18. Norén GN, Edwards R. Opportunities and challenges of adverse drug reaction surveillance in electronic patient records. *Pharmacovigil Rev.* 2010;4(1):17–20.