Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Improving reporting of adverse drug reactions: Systematic review

This article was published in the following Dove Press journal:
Clinical Epidemiology
27 May 2009
Number of times this article has been viewed

Mariam Molokhia[1]
Shivani Tanna[2]
Derek Bell[3]

[1]Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK; [2]Department of Primary Care and Social Medicine, Imperial College, London, UK; [3]Division of Medicine, Imperial College London, Chelsea and Westminster Hospital, London, UK

Background: Adverse drug reactions (ADRs) are a significant cause of morbidity and mortality, with many being identified post-marketing. Improvement in current ADR reporting, including utility of underused or innovative methods, is crucial to improve patient safety and public health.

Objectives: To evaluate methods to improve ADR reporting via a systematic literature review.

Methods: Data sources were Medline, Embase, Cochrane Library and National Library for health searches on ADR reporting (January 1997 to August 2007) including cross-referenced articles. Twenty-four out of 260 eligible studies were identified and critically assessed. Studies were grouped as follows: i) spontaneous reporting (11); ii) medical chart/note review (2); iii) patient interviews/questionnaires (3); and iv) combination methods including computer-assisted methods (8).

Results: Using computerized monitoring systems (CMS) to generate signals associated with changes in laboratory results with other methods can improve ADR reporting. Educational interventions combined with reminders and/or prescription card reports can improve hospital-based ADR reporting, and showed short to medium term improvement.

Conclusions: The use of electronic health data combined with other methods for ADR reporting can improve efficiency and accuracy for detecting ADRs and can be extended to other health care settings. Although methods with educational intervention appear to be effective, few studies have reviewed long-term effects to assess if the improvements can be sustained.

Keywords: adverse drug reaction, reporting, ADR

Introduction

The World Health Organization (WHO) defines adverse drug reactions (ADRs) as "a response to a drug which is noxious and unintended, and which occurs at doses normally used in man for the prophylaxis, diagnosis, or therapy of disease, or for the modification of physiological function".[1] ADRs lead to considerable morbidity and mortality worldwide. A study examining hospital-based admissions for ADRs estimated ADR-related hospital admissions to be as high as 6.5%,[2] although certain patient groups (children and obstetric and gynecology related admissions) were excluded. ADRs were shown to be directly or indirectly associated with as many as 100,000 deaths per year in the USA.[3] In England, hospital episode statistics (HES) data showed that between 1998 and 2005 there were 447,071 ADRs representing 0.50% of total hospital episodes; over this period the number of ADRs increased by 45%.[4] Hence, ADRs have a major impact on public health, reducing patients' quality of life

Correspondence: Mariam Molokhia
Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, WC1E 7HT, UK
Tel +44 20 7927 2633
Email mariam.molokhia@lshtm.ac.uk

and increasing mortality and morbidity, whilst at the same time imposing a considerable financial burden on health care systems. A recent National Confidential Enquiry into Patient Outcome and Death (NCEPOD) report provided an analysis of 649 cancer patients who died within 30 days of receiving chemotherapy in UK hospitals,[5] which further highlights the profile of serious ADR reporting.

Although some ADRs become apparent during clinical trials, many ADRs are not identified until the post-marketing stage. The sample sizes in most clinical trials are often too low for the detection of rare ADRs, and are unable to detect ADRs with long latency periods. Most systems utilized for post-marketing drug surveillance rely primarily on spontaneous reporting. Examples of such systems include the Yellow Card scheme in the UK, which is managed by the Medicines and Healthcare products Regulatory Agency (MHRA), Medwatch in the US, managed by the Food and Drug Administration (FDA) and the Uppsala Monitoring Centre in Sweden, which is the base for the WHO program for International Drug Monitoring. Typically, signals are identified using disproportionality measures such as the reporting odds ratio (ROR) which can be used to estimate relative risk, or the proportional reporting ratio (PRR): proportion of spontaneous reports for a selected drug related to a specific adverse outcome, divided by the corresponding proportion for all or several other drugs.

There are several limitations with spontaneous reporting databases. Although spontaneous reporting provides data on a broad spectrum of patients, the absence of a control group and lack of denominator data mean rate of ADRs cannot be accurately calculated.[6,7] A significant problem lies with under-reporting and biases inherent in clinicians' decisions to report ADRs. Various factors have been found to contribute to under-reporting, with examples including: i) lack of awareness on the purpose of ADR monitoring and how to report ADRs; ii) limited access to yellow cards (although online reporting is now available); iii) uncertainty of reactions being caused by drugs; iv) considering ADRs to be too common or trivial to report; and v) time constraints on clinicians.[8,9] Bias due to media coverage of selected ADRs, and lack of verification of reported diagnoses further limits the data.

New opportunities involve computer-based surveillance methods particularly for hematological, renal, or hepatotoxic ADRs from either secondary or primary care. The key advantages are systematic flagging of abnormal results which warrant further investigation by the responsible physician. This article discusses relative strengths and weaknesses of studies using computerized monitoring systems (CMS) in more detail later.

Methods

We evaluated methods to improve ADR reporting via a systematic literature review. We searched Medline, the Cochrane Library, Embase and the National Library of Health for studies examining methods of improving the reporting of adverse drug reactions published between January 1997 and August 2007. The search strategy used the key MeSH terms and keywords: "adverse drug reactions" OR "adverse drug reaction reporting" OR "adverse drug reaction reporting systems". The search was limited to English language studies. The types of articles included clinical trials, meta-analyses, randomized controlled trials, comparative studies and reviews (Figure 1).

Two reviewers independently selected articles for inclusion from those retrieved by our search. We assessed the studies that met our eligibility criteria as described below. Specifically, for trials this included randomization, allocation concealment, intention to treat analysis and completeness of follow up. Other than trial data we also included studies that met the inclusion criteria and examined methods to improve ADR reports rather than simply reporting prevalence or incidence of ADRs. These papers included data from a number of sources and study groups. We included papers in which the interventions used were described in detail with defined outcome measures.

This search strategy resulted in the identification of 260 studies. Of these 260 studies, 217 were excluded by reviewing the abstract or the title (the majority of these articles were descriptive of incidence or prevalence of ADRs), resulting in 43 potentially eligible studies. Nineteen studies were eligible for inclusion after reviewing the full articles; five further studies were identified from the reference sections of these studies.

Study selection

Inclusion criteria were studies that: i) evaluated or compared methods of ADR reporting; ii) were Europe- or North America-based. Exclusion criteria were studies that: i) were based on unverified patient self reports; ii) evaluated medication errors or prescribing errors specifically; iii) aimed at calculating incidence of ADRs; iv) aimed at identifying specific ADRs; v) aimed at ways of reducing the occurrence of ADRs; vi) described obstacles to ADR reporting; and vii) were purely theoretical papers.

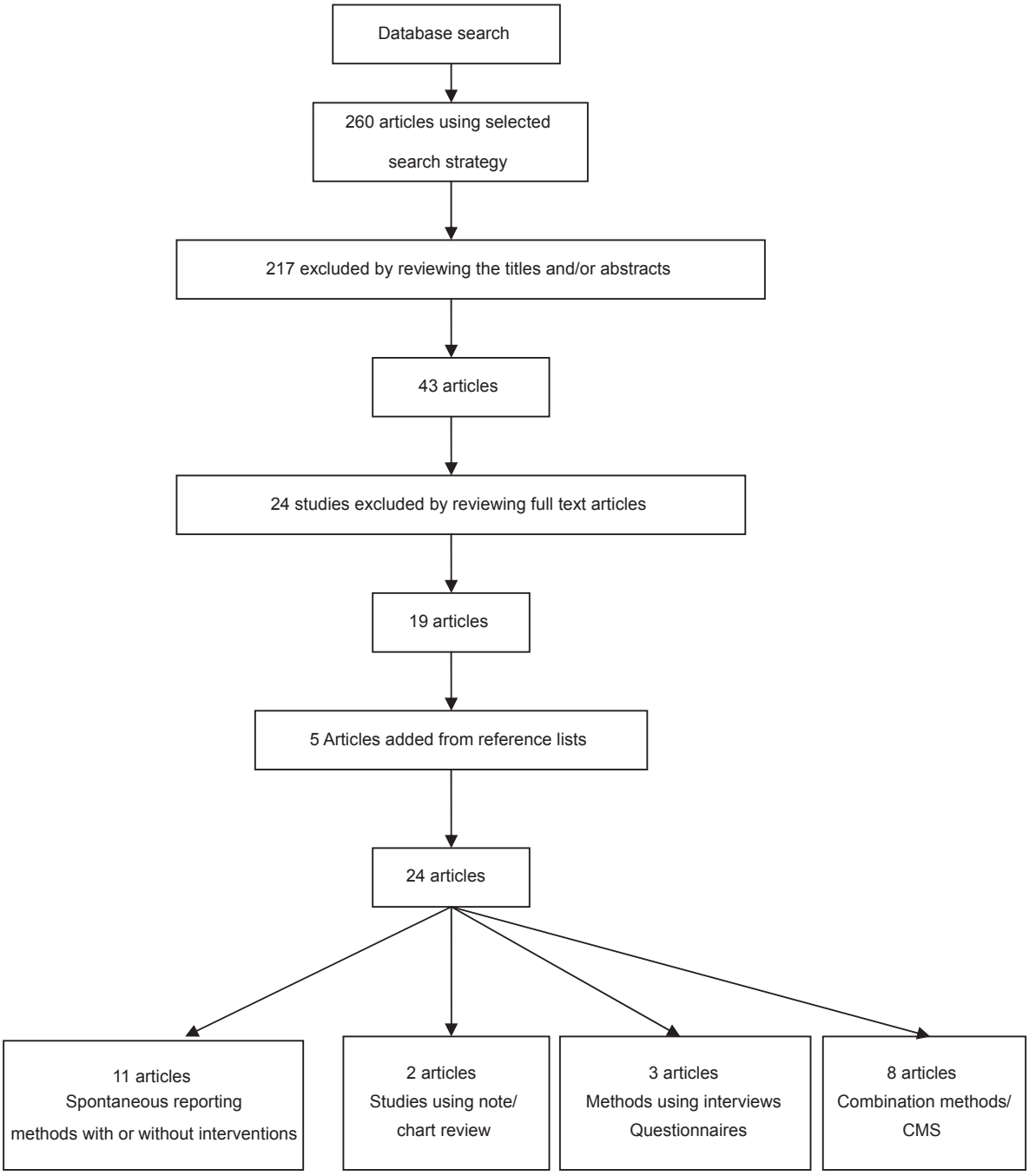

Figure 1 Article selection process for systematic review on ADR reporting.
Abbreviations: ADR, adverse drug reactions; CMS, computerized monitoring systems.

Studies were grouped by method of ADR detection. These were: i) spontaneous reporting and routinely collected data, with or without intervention; ii) note-based and chart reviews; iii) interviews, questionnaires, or observers; and iv) combined methods including computer-assisted methods. For each study, information was collected on: the country in which it was carried out, study type, study period, sample size, the health care professional (HCP) being targeted, the mechanism which was being tested, the outcome measures, and the main conclusions with statistical analysis.

Results

Of the 24 articles selected, four articles were related to trials. For the trials reviewed (including meta-analysis, a total of eleven trials), outcomes included ADR reporting rates, quality measures of reporting,[10–12] and relative

Table 1 Outcomes of studies conducted to evaluate ADR reporting using spontaneous reporting and routinely collected data, with or without intervention

Study	Country	Study type	Study period	Sample size	HCP targeted	Outcome	Statistical test						Conclusion
Disproportionality methods													
Van Puijenbroek[17]	Nether-lands	Comparative study	All reports received from 1982 to Jan 2000	26,555 reports	Academic, Pharmacovigilance	Comparing different measures of disproportionality using PPV (positive predictive value)	PPV	ROR (1.96 SE > 1)	PRR (1.96 SE > 1)	Yules Q (1.96 SE > 1)	χ^2 ($p < 0.05$)	Poisson ($p < 0.05$)	Different measures were broadly comparable only when four or more cases per combination were used
							≥2 reports	0.41	0.44	0.38	0.46	0.34	
							≥3 reports	0.70	0.73	0.67	0.73	0.61	
							≥4 reports	0.83	0.85	0.80	0.82	0.75	
							≥6 reports	0.91	0.90	0.89	0.88	0.84	
Educational Interventions													
McGettigan[18]	Ireland	Interven-tional	9 months (Year not specified)	Doctors 118 questionnaires (65% response)	Physicians	Questionnaire addressing reporting of ADRs and intervention by verbal and written reminders to doctors about ADR reporting and yellow cards placement in drug chart and improving availability in wards/clinics etc	Baseline reports: average 4–6 reports per 3 months 1st 3 months: 24 reports (with reminders and increased availability of yellow cards) 2nd 3 months: 14 reports (after cessation of reminders) 3rd 3 months: similar number of reports to baseline (after removing yellow cards from drug charts)						Only 45% doctors had ever reported an ADR. A four-fold increase in reporting rates was demonstrated but rates reverted to baseline once the intervention was removed
Clarkson[20]	UK	Interven-tional study	12 months (1998 to 1999)	95 reports; 171 ADRs	Physicians	Comparison of rates of reporting before and after introduction of a pilot Paediatric Regional Monitoring Centre (PRMC) and after monthly reminder letters/ yellow cards sent out	12 months before intervention 40 yellow cards received from Trent region 12 months after intervention 113 yellow cards received (95 sent to PRMC and 18 sent direct to MCA, now MHRA) from the Trent region 18/95 (19%) reports and 26/171 (15%) ADRs considered medically significant 24/95 reports contained a drug used "off-label"						Demonstrated an increase in the number of ADR reports using monthly reporting cards; however only 19% reports considered "medically significant" suggesting variable quality 5/18 (28%) reports refer to a dug specifically mentioned in the letter
Castel[19]	Spain	Interven-tional study	Reports from Jan 1983 to Oct 1995	6240 reports	Physicians	Comparison of rates of reporting before and after sending quarterly adverse drug reaction bulletins to all physicians and after introducing yellow cards in prescription pads	Initially, mean monthly reporting rate 34.4 (SD 14.1; n = 106 months) After bulletin sent out (quarterly), mean reporting rate increase to 53.9 (SD = 14.4; n = 48 months) Monthly mean increase in monthly reporting rate with yellow card in prescription pads 19.8 (57.5% increase)						Apparent increase in reporting rates however no control group and subject to potential confounding; only 22% prescribers used prescription pad yellow cards
Morrison-Griffiths[13]	UK	Intervention comparative study	21 months (N/A)	177 reports	Nurses	Nurses given information packs and attended a one-hour teaching session on ADR reporting Assessment of quality of nurse reports compared to doctors	Proportion of reports judged as appropriate according to regulatory criteria Nurses 137/177 (77%); Doctors 676/984 (69%) (95% CI for the difference between proportions 1.4–15.0; p = 0.02) Proportion of reports classified as either probable or possible (Naranjo criteria) Nurses 97% Doctors 98%						Nurses were able to report ADRs appropriately but unclear if this was a result of the educational intervention

Descriptive studies

Barrow[22]	UK	Descriptive study	5 years (1996 to 2000)	Not specified	Physicians	Comparison of HES data for "drug-induced" disorders with spontaneous ADR data for the same reactions during the period 1996–2000 Most ADRs more likely to be recorded in the HES database	Year	1996	1997	1998	1999	2000	ADRs in general more likely to be recorded in HES database (apart from ototoxic hearing loss and hemolytic anemia), where patient episode reporting is mandatory; however information is usually completed by junior staff members
							ADRs with ratio < 100% (more likely recorded in the HES database)	23	25	24	26	25	
							ADRs with ratio > 100% (more likely recorded in yellow card database)	2	2	3	2	1	
Bousquet[23]	France	Comparative study	Reports from Jan 2001 to Jan 2003	42,284 reports	Pharmacovigilance	Comparing signal detection using traditional MedDRA (medical dictionary hierarchy with new approaches which integrated terminological reasoning and Bayesian measures *MedDRA* – Without terminological reasoning *MedDRA TR* – Terminological reasoning by subsumption inside the MedDRA original hierarchy *Ontology TR* – Terminological reasoning by subsumption inside the ontological structure *Ontology TR & AM* – Terminological reasoning by subsumption and approximate matching inside the ontological structure	Significant differences in mean number of occurrences between the MedDRA/MedDRA RT, MedDRA RT/Ontology RT and Ontology RT/ontology & AM with each signal detection technique ($p < 0.001$) Highest mean number of occurrences with Ontology RT & AM (3.63) Lowest mean number of occurrences with MedDRA (2.66)						Ontology terminological reasoning and approximate matching gave highest yield but methods were unvalidated from other SR systems
Macedo[24]	Portugal	Comparative study	Jan 2001	200 reports (131 from GPs and 69 from pharmacists)	Pharmacovigilance	Comparison of published algorithms to evaluate causality in reported ADRs with the currently used WHO GI (general introspection) method Observed agreements when assessing causality between the GI methods and decisional algorithms	Average agreement between GI methods and other decisional algorithms 47% (range 21%–56%) Average kappa score 0.26						High probability of agreements being due to chance as kappa score was only 0.26. Confounding variables not considered could have accounted for the low concordance

(*Continued*)

Table 1 (*Continued*)

Study	Country	Study type	Study period	Sample size	HCP targeted	Outcome	Statistical test	Conclusion
Banks[16]	US	Comparative study	Data collected until 2005	14800 vaccine-COSTART pairs	Pharmacovigilance	Disproportionality Comparing data mining methods – proportional reporting ratio (PRR), screened PRR (SPRR), EMPIRICAL Bayes geometric mean (EBGM) and the lower bound of the EBGM's 90% confidence interval (EBO5)	COSTARTs: Coding Symbols for aThesaurus of Adverse Reaction Terms describe signs, symptoms, and diagnoses	Number of vaccine-COSTART pairs that ranked in the top 100 by each of the methods had substantial variation (42 to 67) and failed to consistently detect previously known vaccinations
RCTs								
Rosebraugh[11]	US	RCT	N/A	85 medical students	4th year Medical students	Intervention with 15 minute lecture on ADR reporting Comparing intervention with non-intervention on quality of reports	Overall quality scores for intervention group were significantly higher than the nonintervention group ($p < 0.004$)	Improvement of ADR report quality but no long term follow up
Figueiras[10]	Portugal	Cluster RCT	Intervention March 2004 to July 2004 Follow up 13–16 months	Intervention group 1388 physicians Control group 5063 physicians	Physicians	Educational intervention with an outreach visit (1 or 2 part presentation) and reminder cards Outcome was change in rate of ADR reporting and duration of the effect of intervention	*Total ADR reporting at start* (*ADRs per 1000 physician-years*) Intervention baseline 7.6 (95% CI 4.0–12.6) Control baseline 11.3 (95% CI 8.9–14.1) P = 0.39 *Total ADR reporting at end of study* Intervention 100.2 (95% CI 85.2–116.4) Control gp 14.5 (95% CI 12.0–18.0) $p < 0.001$ Adjusted increase in rate of total ADR reports attributable to intervention = 90.19 (95% CI 54.51–125.87); RR 10.23 95% CI 3.81–27.51 *Total ADR reporting in first 4 month interval* Intervention 205.2; Control 13.1P < 0.001 *Total ADR reporting in fourth 4 month interval* Intervention 55.3; Control 15.1 P < 0.07	Interventions resulted in a 10-fold increase in reporting rates; but this effect was attenuated with time. Report quality was also improved with more serious, high causality, unexpected and new-drug related ADRs

frequencies and rank order of ADR compared with different reporting methods.[7] Two trials used interventions such as an educational lecture.[10,11] A review of randomized controlled trials (RCTs) compared spontaneous reporting versus solicited collection methods for ADR reports.[12] A meta-analysis used six double-blind, placebo-controlled RCTs to assess frequencies of amiodarone-related ADRs.[7] Details of intention to treat analyses were described,[7,10] however, for the others it was unclear whether an intention to treat analysis was used.

From the individual trials, one explicitly stated a sample size calculation to determine numbers of participants to recruit.[10] Exclusion criteria were given in two papers (seven trials).[7,10]

Discussion

Spontaneous reporting methods with or without interventions

Several studies examined interventional or comparative SR with other data; however there were only two RCTs.[10,11] The most commonly used interventions were verbal and written reminders; supplemented with yellow card placement. Although most intervention studies demonstrated an improvement in ADR reporting, in the largest RCT, follow up was limited to 16 months and so longer term outcomes could not be assessed. Other trained health professionals (nurses and medical students) showed competency in ADR reporting with appropriate interventions,[11,13] but in general were based on small numbers with limited follow up time. Many studies also lacked control groups. There is potential to increase the scope of SR with online yellow card reporting now available both for health care professionals and patients in the UK and US, although the reports may be of variable quality, with little formal evaluation.

Measures of disproportionality

Currently, drug regulators use numerator dependent methods such as ROR, PRR, and Bayesian probability based data mining tools as denominator data is often unavailable.[6] Assuming statistical independence between drug and event, they assess the deviation of the reporting frequency from the expected. However recent evaluation of the WHO Bayesian approach showed good overall sensitivity but rather low specificity, limited by the small number of occurrences of each drug-event association in the database,[14,15] and may yield false positives.[6,16] van Puijenbroek[17] examined concordance of various measures of disproportionality using SR databases including a variety of measures compared to the Bayesian Confidence Propagation Neural Network (BCPNN) analysis. The BCPNN is a probability based approach using linked and new data fields used by the Uppsala Monitoring Centre as a reference measure, yielding an "information component (IC)". The different measures were all highly sensitive but had low specificities compared to the BCPNN analysis, especially with low numbers of reports. Although they were all broadly comparable when four or more reports per combination were present, this was only applicable in 11.2% of Dutch SR cases (unvalidated).[17]

Educational interventions

McGettigan reviewed the effect of increasing availability of yellow cards on wards in Dublin teaching hospitals, Ireland by sending them to prescribers (430 participating doctors) and placing them in drug charts. In addition, verbal and written reminders to doctors about ADR reporting were employed. A four-fold increase in reporting rates was demonstrated, but rates reverted to baseline once the intervention was removed.[18] The increased reporting could be have been attributed to the questionnaire that was sent out about ADR reporting as opposed to the reminders and increased availability of yellow cards. This may have had a confounding effect, as well as seasonal variation bias.

A similar retrospective time series study with 30,000 participating physicians between 1983 and 1995 by Castel showed a 58% increase in mean monthly reporting rates by sending quarterly bulletins about ADR reporting and improving yellow card availability; which declined to 36% and 18% in the second and third month respectively, but lacked a control group. The bulletins were initially sent out quarterly by the end of 1985 however subject to potential confounding, as mailings later varied to monthly or bimonthly.[19] Other confounding factors which were not assessed included the number of new drugs marketed, seasonal effects, population changes, potential reporter training, and variable bulletin content. The inclusion of yellow cards in prescription pads may not have been effective as only 22% of prescribers in the catchment area actually used these pads.

Clarkson compared numbers of yellow cards submitted before and after monthly reminder letters, in addition to spare yellow cards being sent to physicians and establishing a pilot Paediatric Regional Monitoring Centre (PRMC) as an extension of the UK's spontaneous ADR reporting scheme. A 2.8-fold increase in the number of yellow cards received was shown over a 12 month period.[20] It may be easier to identify ADRs in children, who tend to have fewer medicines prescribed, however only 19% of the reports had ADRs that

Table 2 Outcomes of studies conducted to evaluate ADR reporting using note/chart review

Study	Country	Study type	Study period	Sample size	HCP targeted	Outcome	Statistical test	Conclusion
Lata[25]	US	Observational study	Reports from 2000 to 2001	542 ADR reports	Nurses Pharmacists	Educational intervention comparing incidence of adverse drug reaction reports from 1999 using nurse case managers (NCM) and ADR reporters Comparing no of reports by staff nurse and case management dept (NCMs and pharmacist)	No of ADR reports per 100 admissions: 1998 2.1 2000 4.5 2001 5.3 % of all ADRs reported (2000–2001) Case management dept 63 (75% serious ADRs) Staff nurses 26 (9% serious ADRs)	Demonstrated increase in number of ADR reports per 100 admissions and particularly serious ADRs although other confounding factors were not considered. The skills for the NCMs were not standardized
Hougland[26]	US	Descriptive study	One year (2001)	3103 acute care inpatients; 1961 random inpatients; 1142 inpatients sampled from discharge records with ≥ flagged ADE code	Physicians	Assessing validity of ICD-9-CM codes for detecting ADRs using retrospective chart review with a structured tool	Overall PPV for a flagged code representing an ADR was 66%; PPV for inpatient ADRs was 16% For inpatient ADRs from random sample, sensitivity of the flagged codes 10% (23/ 224) and specificity 97% (1689/1737) For ADRs causing hospital admission from random sample, sensitivity of the flagged codes 55% (32/58) and specificity 97% (1855/1903) Within each group wide variation in PPV of various ADR classes	Flagged codes can aid detection of ADRs causing hospital admission. Sensitivity and PPV using flagged codes was much lower for inpatient ADRs than for those related to hospital admission. However ICD-9CM was unable to differentiate events that happened before or during hospital admission

were considered to be medically significant, suggesting variable quality of the reports. The existence of a PRMC may not be generalizable to general practitioners and no control group was examined.

Morrison-Griffiths evaluated the quality of three different nursing group ADR reports, compared to doctors in a UK setting. Nurses were given information packs and attended one-hour teaching sessions on ADR reporting.[13] Seventy-seven percent of nurses' reports were judged as "appropriate" according to the criteria established by the regulatory authority compared to 69% of doctors' reports, and comparable with doctors for causality 97% and 98% respectively using Naranjo criteria;[21] (there are more recent recommendations on causality reporting from the WHO). However it is unclear whether the results are due to effects of the educational intervention they received, as there was no control group. Nurses only submitted reports with agreement of the responsible clinician, although only 3/373 (1%) cards were blocked in this way.

Barrow carried out a descriptive study to compare HES data for ADRs leading to admissions with spontaneously reported ADRs using UK yellow card data between 1996 and 2000. MedDRA codes were identified for the relevant HES discharge diagnoses coded using ICD-10 codes. In particular, HES was more likely to identify drug-induced nephropathy, drug-induced aplastic anemia, dystonia, and Parkinsonism than yellow card data.[22] The study showed that more ADRs are more likely to be recorded in the HES database compared to the yellow card database (apart from ototoxic hearing loss and hemolytic anemia), which may reflect seriousness of ADR (for example warfarin-related bleeding). This may be because it is compulsory to report patient episodes in the HES database; however the information recorded is often completed by the most junior medical team member or by administrative staff. Furthermore, ICD-10 codes used for diagnostic coding in the HES database often lack specificity. HES data contains details of all admissions to NHS hospitals in England, but the yellow card database is used throughout the UK. The matching process for both datasets in terms of time, place, and codes was also imperfect as terminology was different. Despite these limitations the HES database has good potential in being utilized as a mechanism for ADR reporting and monitoring.

Bousquet compared a new approach of signal detection which integrated hierarchical groups and Bayesian measures for signal generation, with the standard method using the MedDRA terms from over 42,000 reports from the French SR database. Five measures of automated signal detection were applied using each method and the mean frequency of drug-adverse effect associations compared. Using these enhanced methods for signal detection, the numbers identified were significantly higher but lacked external validation. The highest mean number of occurrences was found combining the ontology terminological reasoning (TR) and approximate matching (AM) methods (3.63), with significant differences in the mean number of occurrences between approaches ($p < 0.001$).[23] Bousquet's study was limited as the signals generated by each method were unvalidated by other SR systems. Therefore, although more signals were generated by applying the new algorithms, there was no distinction between irrelevant "noise" and relevant signals.

Causality assessment of ADRs

Macedo compared different algorithms by assessing agreement between them, using the WHO "GI" standard method with decisional algorithms using general practitioners (GPs) and pharmacists. The kappa index of reliability was used to quantify the extent to which the observed proportion of agreements exceeded the proportion of agreements expected by chance alone. This was 0.26, suggesting a high probability of agreements being due to chance, particularly where the number of reports was below four.[24] Confounding factors not considered when evaluating algorithms could have accounted for the low concordance.

Figueiras conducted a cluster RCT over a 16-month period in Northern Portugal, covering all National Health System physicians which showed a significant (>10-fold) increase in reporting rates following education intervention and reminder cards, however the effect only remained significant for a year. At the end of the study, there was a significant increase in reports submitted by the intervention group ($p < 0.001$). After 13 months, this difference became less significant ($p = 0.07$) indicating that the duration of the effect of the intervention is time-dependent. Intervention also improved report quality by increasing the reporting rate for serious, high causality, unexpected and new drug-related ADRs.[10] The risk of cross-contamination between groups was minimized by using a cluster-based distribution, and by adjusting for unequally distributed variables. However the study was limited as only 47.2% of physicians assigned to the intervention group actually attended the lecture, suggesting a potentially greater effect with higher attendance, and long term benefits were not assessed. Rosebraugh also noted improvement quality of reports by fourth year medical students following a 15 minute lecture on ADR reporting, although there was no long term follow-up.[11]

Table 3 Outcomes of studies conducted to evaluate ADR reporting using interviews, questionnaires, or observers

Study	Country	Study type	Study period	Sample size	HCP targeted	Outcome	Statistical test	Conclusion
Somers[29]	Belgium	Prospective comparative study	9 months (Mar to Nov 1999)	168 patients and 56 patient interviews	Physicians Nurses Pharmacists	Comparing spontaneous reporting by physicians and nurses with patient interviews by pharmacists as methods for detecting ADRs	Spontaneous ADR reports (12): Nurses 4/168 patients (2.4%) Physicians 8/168 patients (4.8%) Causality of ADRs (n = 12) 75% probable; 25% possible Proportion of ADRs resulting in intervention 92% Patient interviews: 32 ADRs reported from 56 patient interviews Causality of ADRs (n = 32) 72% probable, 28% possible Proportion of ADRs resulting in intervention 59%	Solicited techniques improved ADR reporting. The attributed causality % was similar for both groups. There were a higher proportion of ADRs resulting in intervention identified from spontaneous reports compared to patient interviews
Aspinall[27]	US	Comparative study	2 months (Jan to Feb 2001)	198 patients	Physicians	Comparison of voluntary reporting of ADRs with telephone interviews to patients and providers in identifying ADRs	Patient and provider interviews identified total 83 ADRs in 51 patients (13% probable, 86% possible, 1% doubtful) Reporting rate 0.5% (1/198) vs 51/198 (26%) patient and provider interviews Majority (51%) were related to cardiovascular medications and 48% resulted in medication change	Telephone interviews identified a far higher proportion of suspected ADRs compared to the passive reporting system. Medical staff were unaware of nearly half the ADRs despite outpatient review within preceding 72 hours possibly due to time constraints
Greenhill[28]	US	Comparative study	Not stated	59 Interviews (95% by physicians and 5% by nurse practitioner	Physicians	Comparing number of adverse events (ADRs) reported by different questioning methods: general inquiry (GI); the drug-specific inquiry (DSI); or a comprehensive body system review (BSR)	Total adverse events 195 50 (26%) ADRs identified at GI (54% were moderate to severe) 16 (8%) additional ADRs identified at DSI (75% were moderate to severe) 129 (66%) additional ADRs identified at BSR. (37% were moderate to severe) Clinically relevant ADRs leading to medical action GI 15 (31%); DSI 2 (12%) BSR 19 (15%)	Systematic elicitation of ADRs organized by body systems, increased the identification of clinically relevant ADRs that may not be detected by general enquiry. However the study had low power, ADRs were not validated and was based on children using psychotropic medications only limiting generalizability

Methods using note/chart review

Several studies have used chart review either alone or in combination with other methods; the main disadvantage is the intensive time resource required.

Lata[25] showed that introducing and educating nursing case managers (NCMs) using chart review and patient interviews led to an increase in ADR reporting rates from 1998–2001, (2.1–5.3 per 100 admissions) although other confounding factors that may have affected this rise were not considered, including the number of new medications licensed and other incentives about ADR reporting. It is impossible to know the number of ADRs missed by the reporting system and therefore the sensitivity. The skills for NCMs were not standardized and might have differed between settings, limiting generalizability of results. This study resulted in increased awareness of radio-contrast related ADRs and policy change regarding warfarin-related ADRs.[25]

Hougland used retrospective chart review to assess the validity of ICD-9-CM codes for detecting ADRs, examining 1142 inpatient charts in a US hospital. The overall PPV for a flagged code representing an ADR was 66%, comparing favorably with existing CMS. However, sensitivity of flagged codes for inpatient ADRs was only 10% (rates were higher at 55% for hospital admission-related ADRs) indicating that the selected codes only detect a minority of inpatient ADRs. The specificity was 97% in both groups. Bias may have resulted by introducing terms such as "poisoning" for certain medication errors within the ICD nomenclature. This might have influenced hospital coders due to liability related to documenting a poisoning event. Codes could potentially be improved by removing those with poor PPVs and adding relevant additional codes determined by reviewers.[26]

Methods using interviews/questionnaires

Overall detailed questioning may help elicit potential ADRs but many studies are small, lack validation and may not be generalizable.[27,28]

Somers compared patient interviews by pharmacists with spontaneous reporting by nurses and physicians to detect ADRs in elderly hospital ward patients in Belgium. Somers demonstrated that solicited patient interviews by pharmacists yielded more ADRs 57% (32/56; 98% classified a probable or possible with 59% of these resulting in an intervention) than spontaneous reporting by nurses and physicians 4.8% (8/168), however a higher proportion of ADRs reported spontaneously resulted in medical intervention (92%). This may indicate that spontaneous reports yield a greater amount of severe ADRs that need medical action compared to reports from interviews classified as severe. This study had some internal validity as all reports were discussed weekly with the physician to assess causality, severity, type and level of intervention. The exclusion of sick and confused patients may have affected the results as the confusion may have been a result of an ADR. Combining both methods would probably be more useful for examining ADR occurrence.[29] Aspinall compared numbers of ADRs by telephone interview on a randomly selected population sample of 198 patients and passive spontaneous voluntary reporting in the US. Although the follow up was only two months; 83 ADR were identified by telephone interview (99% classified as probable or possible), compared to just a single spontaneous report. Aspinall demonstrated a higher ADR reporting rate using patient telephone interviews compared to passive voluntary spontaneous reporting. This study was limited as it was based at a single academic tertiary care hospital and only included one reviewer.[27] Medical staff were unaware of nearly half of the ADRs that patients suspected, despite having reviewed the patient in an outpatient appointment within the previous 72 hours, suggesting ADRs could be missed, possibly due to time constraints and inadequate medication review.

Greenhill compared different questioning methods by US physicians to identify ADRs in children receiving psychotropic medications. The general inquiry (GI) was used initially, sequentially followed by the drug-specific inquiry (DSI) and the comprehensive body system review (BSR). 195 ADRs were identified during 59 patient interviews and the study showed systematic elicitation of ADRs, organized by body systems, increased the identification of clinically relevant ADRs that may not be detected by general inquiry, although statistical differences were not assessed. However the ADRs identified were not validated, and questioning order effects not considered. The sample size was also small affecting the study's power and based on children using psychotropic medications only, limiting generalizability.[28]

Combination methods

Computerized monitoring systems

CMS are not currently widely used because of lack of experience, uncertainty about efficiency, and lack of appropriate technology. There were eight studies selected which used combination methods; five of which include the use of CMS. Although these studies showed a low predictive value for signals generated by the CMS, with refinement of the signals,

Table 4 Outcomes of studies conducted to evaluate ADR reporting using combined methods of detection and computer assisted methods

Study	Country	Study type	Study period	Sample size	HCP targeted	Outcome	Statistical test	Conclusion
Jha[30]		Prospective cohort study	8 months (Oct 1994 to May 1995)	617 ADRs; 21,964 patient days	Physicians, nurses and pharmacists	Comparison of adverse drug events detected by computer-based monitoring program, chart review and stimulated voluntary reporting	ADRs identified by: Computer-based monitor 275/617 (45%) Chart review 398/617 (65%) Voluntary reporting 23/617 (4%) Severe ADRs 281/617 identified by: Computer-based monitor 139/281 (49%) Chart review 169/281 (60%) Voluntary reporting 11/281 (4%)	Computer based methods identified fewer ADRs than chart review but more than voluntary reporting. The PPV of computer generated alerts was 16% during the first 8 weeks of the study, but after rule modification this rose to 23%. Types of events captured by computer monitoring (quantitative result based) were substantially different to those captured by chart review (symptom based)
Dormann[31]	Germany	Prospective cohort study	7 months (June to Dec 1997)	379 patients; 46 ADRs	Physicians	Comparing a computer-based adverse drug reaction monitoring system (CMS) with stimulated spontaneous reporting (SSR)	Total ADRs detected 46 (21 regarded as possible and 25 as probable) in 45 patients Overall PPV of automatically generated laboratory signals 13% Relative sensitivity of CMS for all ADRs 74% Relative specificity of CMS for all ADRS 75% Relative sensitivity of SSR for all ADRs 37% Relative specificity of SSR for all ADRs 98%	Although the sensitivity of CMS in this study was 74%, PPV was low at 13%. As there were no denominator data only relative sensitivity could be determined. ADR signals indicating hematological or drug concentration pathology had higher PPVs and were more frequently associated with an ADR than other signals
Dormann[32]	Germany	Prospective pharmacoepidemiological survey	7 months (Sept 2000 to March 2001)	377 patients; 474 admissions	Physicians/ pharmacists	Using a computerized monitoring system (CMS) to detect ADRs by laboratory signals and compare to prospective chart review New ALS was the 1st value of a lab test outside the normal defined range in the study Delta ALS was a new value of a lab test which differed significantly from the previous value	PPV of alerts categorized as new ALS was 25% overall and varied between 13% (allergy) and a maximum of 40% (electrolytes) PPV of alerts categorized as delta ALSs was 32% overall (range 18% (alkaline phosphatase) and 67% (serum electrolytes). New ALS: Sensitivity 91% Specificity 23% Delta ALS: Sensitivity 40% Specificity 76%	Introducing signals generated by more significant differences in laboratory values from the normal range improved the overall PPV of alerts. However these were not as sensitive as using alerts generated using the first value of a laboratory test outside the normal range However, the specificity of the new ALS was lower that the delta ALS. The value of PPV was highest for delta ALS regarding sodium and potassium levels (67%) so this could be a useful method of detecting ADR induced electrolyte disturbances

Hope[33]	US	Interventional study	5 months (Jan to May 2001)	Indianapolis 22 clinics (tiered) Boston 11 clinics (pharmacist)	Pharmacists	Tiered approach of identifying ADRs compared to traditional pharmacist based approach Frequency of adverse drug events and medication errors and the positive predictive value (PPV) of a signal for ADRs		No of ADRs Boston 242 Indianapolis 535	PPV of a signal for ADRs 10.2% 9.6% p-value 0.36	Traditional methods of ADR detection reliant on chart review by pharmacists produced ADR signals with low predictive values, and there is little difference between a tiered vs pharmacist approach. There were a number of possible confounding factors between the sites which were not taken into account
Haffner[34]	Germany	Prospective cohort study pediatric general and intensive care admissions	5 months (Feb to June 2001)	52 beds	Physicians	Comparison of ADR reporting using intensified surveillance (chart review) and computer-assisted screening of pathological laboratory parameters as a signal for a potential ADR	Intensified surveillance detected 101 ADRs in 11.9% of patients Computer-assisted screening detected 45 ADRs in 5.7% of patients Sensitivity: Intensified surveillance 67.2% Computer-assisted screening 44.8% Specificity Intensified surveillance not specified Computer-assisted screening 72.8% 25% ADR-related drugs were used for "off-label" indications			PPV of computer-assisted screening signals was 18.6% and detected ADRs associated with hepatic or hematological consequences compared to intensified surveillance which identified gastrointestinal, dermatological and neurological ADRs. Intensified surveillance detected more and more severe ADRs than CMS. Implementation of additional clinical information like admission diagnosis, medication orders, dose variations, diagnostic findings and clinical events could help improve computerized detection
Neubert[35]	Germany	Prospective pharmacoepidemiological survey	6 months (Year not specified)	396 pediatric isolation patients	Physicians	ADRs were identified by intensive chart review Evaluating computerized monitoring system (CMS) based on laboratory test results for detection of ADRs in addition to spontaneous reporting by the treating physician	73 ADRs identified in 439 admissions (52/396 patients) For all ADR signals: Sensitivity 90.3% Specificity 19.6% With alerts for relative laboratory result changes only (DELTA); sensitivity decreased 50% and specificity increased to 75.9% DELTA alerts were significantly more frequent indicators of ADRs (29.7%) than alerts based on static changes (12.5%), $p < 0.05$			Implementing a CMS can be used as an adjunct for ADR detection by the treating physician. CMS are selective for certain ADRs and dependent on physician initiated request DELTA CMS alerts may be useful for identifying trends in laboratory values, however specificity remains low suggesting many false signals may be generated

(*Continued*)

Table 4 (*Continued*)

Study	Country	Study type	Study period	Sample size	HCP targeted	Outcome	Statistical test	Conclusion
Loke[7]	UK	1) Meta-analysis of clinical trials	Studies published between 1997 and 2002	6 RCTs with data on 4000 patients	Academic	Relative frequencies and rank order of ADR relative frequencies compared using three datasets Comparison of rank order of ADRs	1) Highest rank order Heart (relative frequency 1.80) Lowest rank order Skin (relative frequency 0.35)	All three methods showed poor concordance with each other possibly due to various constraints and biases of each method. In the RCTs only 2000 patients were treated with amiodarone which would reduce likelihood of detecting rare ADRs. The value of case reports is limited due to established reporting and publication bias but detail could help supplement spontaneous case reporting in hypothesis generation
		2) Case reports published in medical journals	Case reports published between 1966 to 2000	357 case reports			2) Highest rank order Respiratory (relative frequency 1.00) Lowest rank order GI (relative frequency 0.02)	
		3) Spontaneous reports sent to the WHO	Reports from 1982 to 2001	640 reports			3) Highest rank order Thyroid (relative frequency 1.70) Lowest rank order Eyes (relative frequency 0.20)	
Wernicke[12]	US	3 RCTs Study 1. children, Study 2. children and adolescents, Study 3. adults	Not stated	653 patients	Physicians	Using solicited (questionnaire) and unsolicited (spontaneous reporting) methods to ascertain ADRs reporting rates with drug and placebo Sp-So index calculated by dividing the spontaneous D/P ratio by solicited D/P ratio D/P = drug/placebo reporting ratio for a selected event	Of the 29 adverse events, 100% had higher rates of reporting using solicited methods The unsolicited method had the greatest ability to distinguish drug from placebo in 22/29 (76%) of the events, however these results were not statistically significant for 20 of the events listed	Although sensitivity might be greater using solicited methods, specificity may be compromised as patients may report insignificant symptoms. The RCTs all used different checklists and there was lack of standardization of symptoms, diagnosis, drug dosage and treatment duration

this may improve. Chart review, although better at identifying ADRs, has the disadvantages of being expensive and time-dependent, with limited scope for intervention. CMS can be prospective, allowing early detection and intervention in drug therapy. It involves less staff and has been shown to detect more ADRs than stimulated spontaneous reporting (SSR). In general methods to improve ADR reporting have been most successful when combined with CMS. Additionally, many software programs exist to implement automated signal detection in post-marketing databases.[6]

Jha carried out an eight-month prospective cohort study in a tertiary care teaching US hospital to compare ADRs detected by a CMS, chart review and stimulated voluntary reporting, and found that computer-based methods identified fewer ADRs (45%) than chart review (65%), but more than voluntary reporting (4%). The PPV of computer-generated alerts was 16% initially, but after later changing rules (9/49), this increased to 23%. This study demonstrated that the types of events captured by computer monitoring are substantially different to those captured by chart review. The advantages of this study are that it was prospective in nature and that the independent reviewers assessing the likelihood of alerts being due to ADRs, were blinded to the data generated by the computer monitor to limit bias.[30] Limitations include the absence of a gold standard for comparison.

Dormann used a prospective cohort study over six months to compare a CMS for automatically generated laboratory signals with SSR in a German university hospital. Both CMS signals and SSRs were validated by the pharmaco-epidemiological team, using chart review. The PPV of automatically generated laboratory signals by the CMS was 13%. The sensitivity and specificity of the CMS was 74% and 75%, respectively. The sensitivity of SSR was 37% but the specificity was higher at 98%.[31]

In this study, due to absence of denominator data, only the relative sensitivity could be determined. Certain ADRs, for example, signals indicating hematological pathology or drug concentration, had higher PPVs (17%–25%) and were more frequently associated with an ADR than other signals. SSR was most effective in detecting clinical symptoms such as gastrointestinal side effects, whereas the CMS more reliably identified ADRs associated with quantitative changes in laboratory values such as nephrotoxicity, hepatotoxicity, and hematological changes.[31] In a later study, Dormann compared a CMS with prospective chart review over six months to detect possible ADRs by laboratory induced abnormalities. In addition, they compared the PPV of different signals known as "new ALS" (automatic laboratory signal; the first value of a laboratory test outside the normal defined range); and "delta ALS" (a new value of a laboratory test which differed significantly from the previous value). The PPV of alerts categorized as "new ALS" varied between 13% for immunoglobulin E (IgE)-related allergy, and a maximum of 40% for change in electrolytes (calcium, potassium, sodium). The overall PPV of alerts categorized as "delta ALS" was higher (18% to 67%) for abnormalities in liver enzymes (ALP) and sodium and potassium levels, respectively. However, these were not as sensitive (40%) as using alerts generated using the 1st value of a laboratory test outside the normal defined range (sensitivity 91%). However, the specificity of alerts generated by the "new ALS" was lower at 23% compared to 76% with the "delta ALS".[32]

Hope compared the use of a tiered approach to a traditional pharmacist based approach for identifying ADRs with patients attending ambulatory care clinics in Indianapolis and Boston. The tiered approach consisted of generating signals from electronic medical records by computer queries, followed by exclusion of some signals by nonclinical data managers who could also add data, subsequently reviewed by nurses and pharmacists. The alternative approach was pharmacist based review. Computer searches were used in both groups for the initial detection. There was no significant difference in PPVs ($p = 0.36$) by either approach, 10.2% (tiered) versus 9.6% (pharmacist), which were both low.[33] The tiered approach (although deemed cost-effective) is a complicated process that may be difficult to enforce and relies on availability, experience, and judgment of pharmacists. Confounding factors include that both sites had different patient populations with unstandardized electronic medical records, which suggests that lack of differences may not be attributable to the different approaches alone.

Haffner used a similar approach in a three-month prospective cohort study in a German teaching hospital to compare ADR reporting using intensified chart review and computer-assisted screening of pathological laboratory parameters. Chart review had a higher sensitivity at 67.2% when compared to computer-assisted screening (sensitivity 44.8%). The specificity for intensified surveillance was not calculated but was 72.8% for computer-assisted screening with a mean PPV of signals of 18.6%. Computer assisted screening identified ADRs associated with hepatic or hematological consequences, whereas intensified surveillance (however no details on training were given) identified more ADRs related to gastrointestinal, dermatological, and neurological symptoms, which was also confirmed in other studies.

Sensitivity was unreliable as it was based on the assumption that all ADRs are picked up by both methods.[34]

Neubert carried out a prospective pharmacoepidemiological pediatric ward-based survey in a University hospital in Germany to evaluate a CMS which uses automatic laboratory signals as alerts for identifying ADRs in addition to spontaneous reporting by the treating physician. Although the sensitivity of the CMS for ADRs was 90.3%, the specificity was only 19.6% indicating that although it identifies most ADRs, it might be generating additional noise. With alerts for relative laboratory result changes only (DELTA); although specificity increased to 75.9%, sensitivity decreased to 50% which was similar to Dormann's findings.[32] This study was limited as laboratory tests were only performed when requested by the treating physician. Routine observation of ADRs by the practicing physician only identified 43% of ADRs, which increased when combining this method with CMS to 74%, however both methods failed to detect 26% of ADRs.[35] These results suggest that implementing a CMS based on laboratory data was ineffective alone and should only be used as an adjunct to observation by the treating physician. In addition, certain ADRs only cause clinical symptoms, for example diarrhea; and these cannot be detected by the CMS. Conversely, some ADRs may only be detected by change in laboratory value, for example drug-induced neutropenia or anemia.

Other combination methods

Loke compared frequencies of ADRs to the antiarrhythmic drug, amiodarone using three different datasets generated from: i) a meta-analysis of RCTs; ii) published case reports; and iii) from spontaneous reports sent to the WHO. The distributions of ADR rank order and relative frequencies were dissimilar among the datasets. For example the highest rank order ADRs were cardiac in the dataset produced by the meta-analysis in contrast to respiratory and thyroid for case reports and spontaneous reports respectively. Results demonstrated poor concordance when comparing amiodarone ADR relative frequencies using datasets generated by a meta-analysis of six RCTs, case reports, and spontaneous reports sent to the WHO.[7] Due to the strict selection criteria applied, only six RCTs were used, which limited the power of the meta-analysis. The strengths of the meta-analysis were that selection and detection bias were minimized as the RCTs were double-blinded and placebo-controlled. However, in the six RCTs, only 2,000 patients were treated with amiodarone which would reduce the likelihood of yielding useful information on rare ADRs. Variation in the types of ADRs reported in the different RCTs could have led to some loss of precision when using a categorized summary of the data. Although 357 case reports fulfilled the inclusion criteria for ADRs to amiodarone, the value of data produced by such reports is limited due to established reporting and publication bias.

Wernicke compared spontaneous and solicited ADR collection methods from three large randomized, double-blind clinical-controlled trials ADRs in children, adolescents, and adults given either a drug or placebo. ADRs were either collected by unsolicited methods (spontaneously) where open-ended questions were asked, or by solicited methods where questionnaires consisting of detailed checklists were administered after spontaneous events were recorded. Solicited methods yielded higher rates of reporting for all 29 ADRs identified. However, in 76% of ADRs, unsolicited methods (SR) had a greater ability to distinguish drug from placebo effects. This suggests that although the sensitivity might be greater using solicited methods, specificity may be compromised as prompting patients to report ADRs might stimulate reporting of insignificant symptoms.[12] Lack of standardization of symptoms, diagnosis, drug, dosage and treatment duration may have also acted as confounding factors and may have affected the perception and reporting of ADRs. In addition, with spontaneous reports, patients might not report events which are transient or embarrassing and have resolved before their visit.

Limitations of our review

Although a systematic search of the literature was undertaken based on our selected criteria, some published and unpublished studies may have been omitted. Most of the studies had very short follow up times, and the longest trial included follow up for 16 months. Therefore we are unable to determine if improvements in ADR reporting are sustained long term.

We were also limited by the data available from the papers and the ability to combine studies due to the heterogeneous nature in the designs. Many studies did not include statistical interpretation of their results. We reported limited meta-analysis for two studies,[7,12] but did not attempt a separate re-analysis. Because of the small number of studies, we were unable to find evidence for publication bias.

Conclusions

There are several limitations with current methods of ADR reporting. Most studies that have looked at the effect of educational intervention to improve spontaneous reporting have shown considerable improvement in rates, although this benefit decreases with time. Increasing the availability of yellow cards on wards as well as encouragement to use

web-based reporting may also improve reporting rates. Questionnaires, chart review and patient interviews can be a valuable source of information on ADRs, but are time and personnel dependent. Other factors that could influence reporting of ADRs are the quality of data in spontaneous reporting databases and the statistical methods employed to detect signals. Improving coding of terms to describe ADRs, to make them more specific and homogeneous, may improve quality of reports. Combining inpatient reviews with hospital-based CMS may better prospectively identify potential ADRs. CMS methods appear to have greater sensitivity than SR (possibly greater for those generated by the first value of a laboratory test outside the normal range); although in the studies we reviewed the positive predictive value was low. The PPV may be better for alerts generated by a laboratory value which differed significantly from the previous value. This could be explained by a number of factors and refinement of the CMS approach such as including filters to exclude nondrug causes may help improve the results.

With health care systems becoming more computerized in both primary and secondary care, there is great potential to explore these systems for ADR detection. For example, the addition of a section for ADRs that occurred during the patient's admission in electronic discharge summaries could improve ADR detection. The computerization of medical health records with prescription data covering several million people can help promote developments of methods for detecting ADRs in clinical practice in the UK and other countries. Ongoing initiatives include the EU-ADR project, a FP7 multinational collaboration for the early detection and investigation of adverse drugs events,[36] using data from a number of European medical databases.

The population of both developed and developing countries are gradually aging, with an increasing proportion of the population comprised of elderly people who are at greatest risk of suffering from chronic diseases. There is also greater pressure on clinicians to treat chronic diseases and their risk factors using evidence-based guidelines. Both these factors are increasing prescribing rates, promoting polypharmacy, and leading to an increased risk of drug interactions and ADRs. Accurate surveillance is an essential first step that must be undertaken to identify ADRs and implementing measures to reduce their public health impact.

Acknowledgments

MM is funded by a post-doctoral fellowship from the National Coordinating Centre for Research Capacity Development, and is an investigator for the FP7 EU-ADR project (http://www.euadr-project.org/). The Department of Primary Care and Social Medicine at Imperial College and Department of Acute Medicine, Division of Medicine is grateful for support from the NIHR Biomedical Research Centre scheme and the NIHR Collaboration for Leadership in Applied Health Research and Care for Northwest London. ST and MM undertook review design and data collection. All authors contributed to critical revisions and writing of the manuscript and approved the final version for publication.

References

1. World Health Organization. *International drug monitoring: The Role of National Centres. Report No: 498.* Geneva, Switzerland: World Health Organization; 1972.
2. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ*. 2004;329(7456):15–19.
3. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*. 1998;279(15):1200–1205.
4. Patel H, Bell D, Molokhia M, Srishanmuganathan J, Patel M, Car J, et al. Trends in hospital admissions for adverse drug reactions in England: analysis of national hospital episode statistics 1998–2005. *BMC Clin Pharmacol*. 2007;7(1):9.
5. National Confidential Enquiry into Patient Outcome and Death [NCEPOD]. For better, for worse? A review of the care of patients who died within 30 days of receiving systemic anti-cancer therapy. 2008. Cited February 3, 2009. Available from: http://www.ncepod.org.uk/2008sact.htm.
6. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. *Drug Saf*. 2003;26(3):159–186.
7. Loke YK, Derry S, Aronson JK. A comparison of three different sources of data in assessing the frequencies of adverse reactions to amiodarone. *Br J Clin Pharmacol*. 2004;57(5):616–621.
8. BMA Science and Education department and the BMA Board of Science. *Reporting Adverse Drug Reactions: A guide for healthcare professionals*. London: British Medical Association; 2006.
9. Stephens MDB. Introduction. In: Talbot J, Waller P, editors. *Stephens' Detection of New Adverse Drug Reactions*. 5th ed. Hoboken, NJ; John Wiley and Sons Ltd.;2005. p. 10–17.
10. Figueiras A, Herdeiro MT, Polonia J, Gestal-Otero JJ. An educational intervention to improve physician reporting of adverse drug reactions: a cluster-randomized controlled trial. *JAMA*. 2006;296(9):1086–1093.
11. Rosebraugh CJ, Tsong Y, Zhou F, et al. Improving the quality of adverse drug reaction reporting by 4th-year medical students. *Pharmacoepidemiol Drug Saf*. 2003;12(2):97–101.
12. Wernicke JF, Faries D, Milton D, Weyrauch K. Detecting treatment emergent adverse events in clinical trials: a comparison of spontaneously reported and solicited collection methods. *Drug Saf*. 2005;28(11):1057–1063.
13. Morrison-Griffiths S, Walley TJ, Park BK, Breckenridge AM, Pirmohamed M. Reporting of adverse drug reactions by nurses. *Lancet*. 2003;361(9366):1347–1348.
14. Lindquist M, Stahl M, Bate A, Edwards IR, Meyboom RH. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. *Drug Saf*. 2000;23(6):533–542.
15. Gould AL. Practical pharmacovigilance analysis strategies. *Pharmacoepidemiol Drug Saf*. 2003;12(7):559–574.
16. Banks D, Woo EJ, Burwen DR, Perucci P, Braun MM, Ball R. Comparing data mining methods on the VAERS database. *Pharmacoepidemiol Drug Saf*. 2005;14(9):601–609.

17. van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf*. 2002;11(1):3–10.
18. McGettigan P, Golden J, Conroy RM, Arthur N, Feely J. Reporting of adverse drug reactions by hospital doctors and the response to intervention. *Br J Clin Pharmacol*. 1997;44(1):98–100.
19. Castel JM, Figueras A, Pedros C, Laporte JR, Capella D. Stimulating adverse drug reaction reporting: effect of a drug safety bulletin and of including yellow cards in prescription pads. *Drug Saf*. 2003;26(14):1049–1055.
20. Clarkson A, Ingleby E, Choonara I, Bryan P, Arlett P. A novel scheme for the reporting of adverse drug reactions. *Arch Dis Child*. 2001;84(4):337–339.
21. Naranjo CA, Busto U, Sellers EM, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther*. 1981;30(2):239–245.
22. Barrow P, Waller P, Wise L. Comparison of hospital episodes with 'drug-induced' disorders and spontaneously reported adverse drug reactions. *Br J Clin Pharmacol*. 2006;61(2):233–237.
23. Bousquet C, Henegar C, Louet AL, Degoulet P, Jaulent MC. Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach. *Int J Med Inform*. 2005;74(7–8):563–571.
24. Macedo AF, Marques FB, Ribeiro CF, Teixeira F. Causality assessment of adverse drug reactions: comparison of the results obtained from published decisional algorithms and from the evaluations of an expert panel. *Pharmacoepidemiol Drug Saf*. 2005;14(12):885–890.
25. Lata PF, Mainhardt M, Johnson CA. Impact of nurse case manager-pharmacist collaboration on adverse-drug-event reporting. *Am J Health Syst Pharm*. 2004;61(5):483–487.
26. Hougland P, Xu W, Pickard S, Masheter C, Williams SD. Performance of International Classification Of Diseases, 9th Revision, Clinical Modification codes as an adverse drug event surveillance system. *Med Care*. 2006;44(7):629–636.
27. Aspinall MB, Whittle J, Aspinall SL, Maher RL, Jr, Good CB. Improving adverse-drug-reaction reporting in ambulatory care clinics at a Veterans Affairs hospital. *Am J Health Syst Pharm*. 2002;59(9):841–845.
28. Greenhill LL, Vitiello B, Fisher P, et al. Comparison of increasingly detailed elicitation methods for the assessment of adverse events in pediatric psychopharmacology. *J Am Acad Child Adolesc Psychiatry*. 2004;43(12):1488–1496.
29. Somers A, Petrovik M, Robays H, Bogaert M. Reporting adverse drug reactions on a geriatric ward: a pilot project. *Eur J Clin Pharmacol*. 2003;58:707–714.
30. Jha AK, Kuperman GJ, Teich JM, et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc*. 1998;5(3):305–314.
31. Dormann H, Muth-Selbach U, Krebs S, et al. Incidence and costs of adverse drug reactions during hospitalisation: computerised monitoring versus stimulated spontaneous reporting. *Drug Saf*. 2000;22(2):161–168.
32. Dormann H, Criegee-Rieck M, Neubert A, et al. Implementation of a computer-assisted monitoring system for the detection of adverse drug reactions in gastroenterology. *Aliment Pharmacol Ther*. 2004;19(3):303–309.
33. Hope C, Overhage JM, Seger A, et al. A tiered approach is more cost effective than traditional pharmacist-based review for classifying computer-detected signals as adverse drug events. *J Biomed Inform*. 2003;36(1–2):92–98.
34. Haffner S, von LN, Wirth S, Thurmann PA. Detecting adverse drug reactions on paediatric wards: intensified surveillance versus computerised screening of laboratory values. *Drug Saf*. 2005;28(5):453–464.
35. Neubert A, Dormann H, Weiss J, et al. Are computerised monitoring systems of value to improve pharmacovigilance in paediatric patients? *Eur J Clin Pharmacol*. 2006;62(11):959–965.
36. EU-ADR: FP7 Multi-national project to integrate electronic health care records for the early detection and investigation of adverse drugs events. 2009. Accessed on January 28, 2009. Available from: http://www.euadr-project.org/.