

Medication errors: how reliable are the severity ratings reported to the national reporting and learning system?

STEVEN D. WILLIAMS^{1,2} AND DARREN M. ASHCROFT^{1,2}

¹Pharmacy Department, University Hospital of South Manchester NHS Foundation Trust, Southmoor Road Wythenshawe, Manchester M23 9LT, UK, and ²School of Pharmacy and Pharmaceutical Sciences, University of Manchester, Manchester M13 9PL, UK

Abstract

Objective. To examine: (1) the reliability of the severity rating scale used by the National Reporting and Learning System (NRLS) in England and Wales for medication errors; and (2) the likelihood of reporting medication errors among healthcare professionals.

Setting. A 900-bed acute university teaching hospital in the North West of England.

Participants. Forty healthcare professionals (10 doctors, 10 nurses, 10 pharmacists and 10 pharmacy technicians).

Methods. Participants were asked to complete a self-administered questionnaire containing nine medication error scenarios on two separate occasions. They were asked to rate the severity of each incident using the NRLS severity rating scale and also the likelihood of reporting the incident via the hospital incident reporting system. The main outcome measures included comparisons of severity ratings and likelihood of reporting by the four health professional groups. Test–retest reliability of the severity ratings was also examined within and between professional groups.

Results. Pharmacists and nurses were significantly more likely to report the errors if they had witnessed them (mean scores 36.3 and 36.2, respectively, compared with 27.9 for doctors, $P < 0.001$). Nurses and pharmacy technicians assigned higher severity ratings for medication errors (mean scores 23.6 and 25, respectively) than pharmacists or doctors (both 19.4). Both within and between healthcare professional groups, there was wide variation in the assignment of medication error severity ratings.

Conclusions. There are marked differences in the severity ratings for medication errors graded against the NRLS severity criteria between different health professional groups and at different time points rated by the same individuals.

Keywords: incident reporting and analysis, drug errors

Introduction

The drive for safer healthcare has never been stronger and it is generally accepted that learning from errors is a vital part in achieving this [1, 2]. In the UK, the National Patient Safety Agency (NPSA) established the National Reporting and Learning System (NRLS) in England and Wales in 2004. The NRLS was designed to co-ordinate the voluntary reporting of patient safety incidents within healthcare and to improve the ability of the National Health Service (NHS) to learn from the analysis of these events. The NPSA defines a patient safety incident as ‘any unintended or unexpected incident that could have or did lead to harm for one or more patients receiving NHS-funded healthcare’ [3]. The NRLS

uses a severity grading system to indicate the scale of harm experienced by the patient as a result of the incident.

Since 2005, the NPSA has published quarterly incident reports from the NRLS; medication errors are the second largest category, accounting for ~8.5% of all incidents reported [4]. Between January 2005 and June 2006, 59 802 medication safety incidents were reported to the NRLS with just over 80% being reported from acute, general or community hospitals [5]. The majority (82.8%) were reported as causing no harm, 16.4% as causing low or moderate harm and a very small proportion reported as causing severe harm (0.8%) or death (0.1%) [5].

Our experience of analysing medication error reports within a large university teaching hospital (that routinely

Address reprint requests to: Mr Steven D. Williams, University Hospital of South Manchester NHS Foundation Trust, Southmoor Road Wythenshawe, Manchester M23 9LT, UK. Tel: 0161 2912113; E-mail: steve.williams@uhsm.nhs.uk

submits data to the NRLS) would suggest that healthcare practitioners often struggle to assign severity ratings to incidents that they have encountered. In particular, inconsistencies in severity ratings have been noted when different healthcare professionals have reported the same medication error, or when 'near miss' events have been detected and differences between actual versus potential level of harm become apparent. This is important as the risk management strategy for the hospital prioritizes more in-depth investigations for 'high' risk events which are determined by the reported severity of the incident and judgments of the likelihood that the event could reoccur in the future.

With additional concerns about the general reluctance of health professionals to report patient safety incidents there is an even greater need to be confident of the reliability of the medication errors actually reported [6–8]. Walshe's review of measuring adverse events in healthcare suggests that consistent inter- and intra-rater reliability of adverse events is elusive with most studies indicating at best moderate to good reliability of adverse event severity ratings [9]. Very few studies have formally examined agreement on rating medication error severity between multiple health professionals [10–12] or considered both the inter- and intra-rater agreements [10–11]. None of the scales used in these studies have been routinely adopted for use throughout the NHS, unlike the NRLS severity rating scale. The aims of this study were, therefore, to examine in four different health professional groups: (1) the reliability of the severity rating scale used by the NRLS for medication errors; and (2) the likelihood of reporting medication errors.

Method

Forty healthcare professionals (10 doctors, 10 nurses, 10 pharmacists and 10 pharmacy technicians) working within a 900-bed acute university teaching hospital in the North West of England agreed to take part in the study. A maximum variation sampling approach was used to ensure that the

participants differed in terms of age and seniority of position across the four professional groups. Following review of the study protocol, the chairman of the local NHS research ethics committee advised that ethical approval was not required for this study.

The participants were asked to complete a self-administered questionnaire containing nine medication error scenarios on two separate occasions, 2 weeks apart. To ensure realism, the contents of the scenarios were based on actual medication errors that had previously been reported to occur in the hospital setting [13]. Piloting of the questionnaire was undertaken with five members of the hospital pharmacy department to ensure that the survey procedures and language were appropriate.

The nine scenarios consisted of three prescribing, three dispensing and three drug administration errors. For each scenario, participants were asked to imagine that they had witnessed the incident occurring in the hospital and then rate the likelihood that they would report the incident via the hospital incident reporting system, using a five point scale ranging from (1) 'very unlikely' to (5) 'very likely'. In addition, they were also asked to rate the severity of each incident using the NRLS severity rating scale ranging from (1) 'no harm' to (5) 'death', as shown in Table 1. An example of a prescribing error scenario used in the survey is shown in Box 1.

Box 1 Example of a scenario used in the questionnaire

Prescribing error scenario.

A 3-year-old child was prescribed IV aciclovir 400 mg for the treatment of chicken pox. The nurse had mentioned to the doctor that the dose appeared large but the doctor instructed the nurse to give the dose as prescribed. The dose had been calculated incorrectly and the patient received ten times the correct dose on one occasion before the error was noticed. As a result of the error, the child experienced an increase in their blood urea and creatinine levels which resolved within 2 days.

Table 1 NRLS patient safety incident severity rating scale

Severity	Description
No harm	Impact prevented: any patient safety incident that had the potential to cause harm but was prevented, resulting in no harm to people receiving NHS-funded care Impact not prevented: any patient safety incident that ran to completion but no harm occurred to people receiving NHS-funded care
Low	Any patient safety incident that required extra observation or minor treatment and caused minimal harm, to one or more persons receiving NHS-funded care
Moderate	Any patient safety incident that resulted in a moderate increase in treatment and which caused significant but not permanent harm, to one or more persons receiving NHS-funded care
Severe	Any patient safety incident that appears to have resulted in permanent harm to one or more persons receiving NHS-funded care
Death	Any patient safety incident that directly resulted in the death of one or more persons receiving NHS funded care

Data analysis

Rating scores for both the likelihood of reporting and severity ratings across all nine scenarios were summed to produce overall scores for each respondent. For each survey round, differences between overall ratings for each health professional group were compared using analysis of variance. The Scheffé correction was used to control for Type I error in *post hoc* comparisons of scores between professional groups. All calculations were performed using SPSS v13.0; a threshold of $P < 0.05$ was set to indicate statistically significant differences in scores.

Bland–Altman plots were used to examine test–retest reliability by comparing the intra-rater agreement of the summed severity scores for each individual between the two survey rounds [14]. In addition to the quantitative survey results, the free text comments from the respondents were also examined. The comments were read independently by both authors and key themes about incident reporting were identified.

Results

The mean age of the participants was 34.1 years (SD 9.38) and the mean duration that they had worked in their professional role was 10.6 years (SD 9.9). Seventy-five percent (30/40) of the participants were female. Table 2 presents the demographic characteristics of the participants by professional group. Thirty-nine (97.5%) participants completed both mailings of the survey, with only one pharmacy technician failing to complete the second version.

Likelihood of reporting errors

Mean ratings for the likelihood of reporting the medication errors in both rounds of the questionnaire indicated that the rank ordering of professional groups most likely to report the errors was: (1) pharmacists, (2) nurses, (3) pharmacy technicians and (4) doctors, as shown in Table 3. There were no statistically significant differences between the four groups for the likelihood of reporting in the first round survey results, but there was a significant difference between groups in likelihood of reporting from the second round of the

Table 3 Likelihood of reporting medication errors by professional group

Health professional	N	Round 1		N	Round 2	
		Mean score	Standard Deviation		Mean score	Standard Deviation
Doctor	10	27.90	5.82	10	27.70	5.17
Nurse	10	36.20	9.80	10	38.04	5.10
Pharmacist	10	36.30	6.80	10	39.20	5.92
Pharmacy Technician	10	33.10	9.21	9	33.67	7.07

Table 4 Overall medication error severity scores^a by professional group

Health professional	N	Round 1		N	Round 2	
		Mean score	Standard Deviation		Mean score	Standard Deviation
Doctor	10	19.40	2.41	10	17.70	3.09
Nurse	10	23.60	5.36	10	24.30	6.63
Pharmacist	10	19.40	4.27	10	18.90	4.56
Pharmacy technician	10	25.00	5.23	9	22.11	4.49

^aFor severity score the higher the number the greater the severity.

questionnaire ($F = 8.223$, $P < 0.001$). *Post hoc* tests found that both pharmacists (mean difference in scores = 11.5, $P = 0.001$) and nurses (mean difference = 10.7, $P = 0.003$) were significantly more likely to report errors than doctors (mean score = 27.7).

Severity rating of errors

Table 4 shows the mean severity ratings by professional group. In both rounds of the survey, nurses and pharmacy technicians assigned higher severity ratings for the medication errors than either pharmacists or doctors. There were statistically significant differences between the four professional groups in both the first- ($F = 4.158$, $P = 0.013$) and the second-round ($F = 3.794$, $P = 0.019$) ratings. *Post hoc* tests on the second round results found that nurses assigned significantly higher severity ratings to the medication errors than doctors (mean difference in scores = 6.6, $P = 0.041$). There were no other statistically significant differences between any of the other professional group comparisons.

Figure 1 presents the Bland–Altman plot used to examine the test–retest reliability of the severity ratings showing differences in overall severity scores between the two survey rounds against the mean score for each participant. Only two individuals achieved the same overall severity score on both rounds (i.e. mean difference = 0), with the difference in severity scores between the two rounds ranging from –8 to

Table 2 Demographic characteristics of participants

Professional group	No.	Female: Male	Mean age (SD)	Mean years of experience
Doctor	10	5:5	30.90 (7.64)	7.15 (7.82)
Nurse	10	9:1	35.80 (12.38)	13.48 (12.40)
Pharmacist	10	7:3	32.70 (8.00)	10.30 (8.73)
Pharmacy technician	10	9:1	36.90 (8.95)	11.50 (10.49)

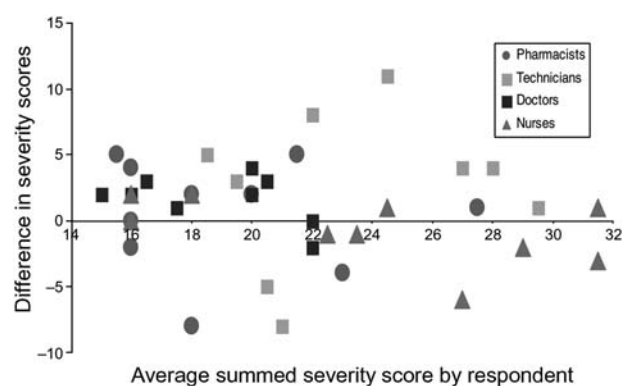


Figure 1 Bland–Altman plot showing variation in severity scores for each participant from repeated ratings.

11. There was also wide variation in the participants' mean severity scores between the two survey rounds with scores ranging from 17.7 to 25 when rating all nine medication error scenarios. Severity ratings from nurses appeared to be most widely spread, with noticeably less variation in severity ratings provided by doctors, as illustrated in Fig. 1.

Health professionals' comments about incident reporting

The analysis of the free text comments provided by the respondents provided additional insights into their attitudes and experiences of reporting medication errors. In particular, perceptions of a prevailing 'blame' culture and the lack of anonymity when reporting errors were identified by several participants as key barriers to error reporting.

Might get more incidents reported if HIRS [Hospital Incident Reporting System] could be filled in without name (Hospital Doctor, female, 30 years old).

The HIRS reporting system is not blame-free, as advertised, but mainly used by those who either have too much free time or are attempting to discredit others... (Hospital Doctor, male 24 years old).

Other respondents highlighted more practical problems when faced with completing an incident reporting form. Within the study hospital, an online incident reporting system was available, but participants felt that at times it was difficult to access the system and that time pressures often meant that errors were not reported.

More incidents would be reported if the system was more user friendly and less time consuming (Hospital Nurse, female, 40 years old).

To be effective the reporting system must be simpler to access and use, and staff must have time to actually be able to report incidents without compromising patient care (Hospital Doctor, male 24 years old).

Several participants described difficulties in assigning the NRLS severity ratings to errors that had occurred. In particular, they felt that the descriptions provided did not necessarily map onto all types of errors that they encountered. Difficulties arose when the error met some of the specified

criteria, but not all the criteria, meaning that judgments were indiscriminate at times.

Use of hospital reporting system can sometimes be difficult to categorise the severity ratings as incidents don't necessarily fulfill all requirements within the desired category (Hospital Nurse, female, 38 years old).

Discussion

Learning from medication safety incidents requires that reporting is reliable and the results from this study suggest that there may well be marked differences in the severity ratings for medication errors reported via the NRLS between different health professional groups and at any given time. In the main, doctors and pharmacists reported medication errors as being less severe than the nurses and pharmacy technicians. This is possibly a reflection of professional training received and the sharing of ideas in practice resulting in similar perceptions of the scenarios. Pharmacists, however, appeared less reliable than their 'front line' medical and nursing colleagues when rating severity for the exact same scenario on a second occasion. The variation of severity ratings was largest for the nurses, perhaps illustrating a greater dependence on personal experience but results for individuals showed acceptable reliability. The pharmacy technicians showed the widest variation and the least reliability in their severity ratings, perhaps reflecting that they have neither been trained nor empowered in their current role to report, and even may not be aware of errors that have occurred if information is not fed back and the outcomes of any errors in practice shared.

Our findings are broadly in line with other studies using different severity scales that have formally examined agreement on medication error severity between multiple health professionals or considered both inter- and intra-rater agreements. In a UK study involving 30 healthcare professionals (10 doctors, 10 nurses and 10 pharmacists), Dean and Barber found 58.7% agreement (calculated using generalizability coefficients) among all the participants, and 78% agreement between the scores given by the same individual on two separate occasions, using a 10 point visual analogue scale for 50 drug administration errors [10].

A study from New Zealand described fair inter-rater agreement (kappa value = 0.34, 95% CI 0.32–0.36) between three raters and moderate to substantial intra-rater agreement (kappa values 0.64, 0.55 and 0.69) using a six point severity scale for 701 paediatric medication errors [11]. However, the raters, two doctors and one pharmacist, were part of a research investigation team. More recently, a US study found substantial overall inter-rater agreement (kappa value = 0.61 {95% CI 0.41–0.81}) for 101 judges using the United States Pharmacopeia's MedMarx nine point severity scale for 27 medication error scenarios [12]. The raters (68.4% pharmacists, 19.4% nurses and 12.2% unclassified) were randomly assigned to evaluate the scenarios either using a simple list of the nine severity definitions, a paper algorithm or a computer-based interactive algorithm to aid severity

classification. The kappa values for the first group, most similar to the methodology used in this study, varied between 0.42 for intervention needed to sustain life to 0.86 for patient death.

Pharmacists and nurses were the professions most likely to report in this study with doctors the least likely to report medication errors if they had witnessed them. Lawton and Parker [15] found that doctors were the least likely to report adverse incidents where the outcome for the patient was bad, with a mean score of 2.97 on a five-point Likert scale compared with mean scores of 4.15 and 3.85 for nurses and midwives, respectively. Evans *et al.* [16] found that if a patient received an erroneous drug treatment that needed corrective treatment 100% of nurses but only 35% of doctors said they would always report the incident.

A key strength of this study was that it compared the severity rating of medication errors, on two different occasions, between four different health professional groups using an incident reporting system routinely used in NHS. Limitations of this study were that it was conducted in a single hospital, and that there is a potential for bias in asking participants if they would report the error in the questionnaire scenarios. The latter may suggest to respondents that consideration to report is expected; therefore this may have resulted in a higher likelihood of reporting than might have been experienced in routine clinical practice. Nonetheless, this study is an important foundation for further work in larger samples to investigate the reliability of reporting medication errors to the NRLS and the possible development of a more reliable severity scale. Equally important is the need to understand the influences on incident reporting by health professionals in order to identify potential interventions to encourage reporting and learning within healthcare.

Acknowledgements

We would like to thank all the health professionals who participated in this research project. Grateful thanks are also extended to Amy Newport and Lucy Bentham for help with distribution of the surveys.

References

1. Kohn LT, Corrigan JM, Donaldson MS. *To Err is Human*. Washington, DC: National Academy Press, 1999.
2. Department of Health. *An Organization with a Memory: a Report from an Expert Working Group on Learning from Adverse Events in the NHS* London: Department of Health, 2000.
3. National Patient Safety Agency. *Seven Steps to Patient Safety - your guide to safer patient care*. London: NPSA, 2004.
4. National Patient Safety Agency. *National Reporting and Learning System Data Summary Issue 7*. London: NPSA, 2008.
5. National Patient Safety Agency. *The fourth report from the Patient Safety Observatory Safety in doses: medication safety incidents in the NHS*. London: NPSA, 2007.
6. Lawton R, Parker D. Barriers to incident reporting in a health-care system. *Qual Saf Health Care* 2006;**15**:196–201.
7. Kingston MJ, Evans SM, Smith BJ *et al.* Attitudes of doctors and nurses towards incident reporting: a qualitative analysis. *Med J Aust* 2004;**329**:12–3.
8. Ashcroft DM, Morecroft C, Parker D *et al.* Likelihood of reporting adverse events in community pharmacy: an experimental study. *Qual Saf Health Care* 2006;**15**:48–52.
9. Walshe K. Adverse events in healthcare: Issues in measurement. *Qual Saf Health Care* 2000;**9**:47–52.
10. Dean BS, Barber ND. A validated, reliable method of scoring the severity of medication errors. *Am J Health Syst Pharm* 1999;**56**:57–62.
11. Kunac DL, Reith DM, Kennedy J *et al.* Inter- and intra rater reliability for classification of medication related events in paediatric inpatients. *Qual Saf Health Care* 2006;**15**:48–52.
12. Forrey RA, Pederson CA, Schneider PJ. Interrater agreement with a standard scheme for classifying medication errors. *Am J Health Syst Pharm* 2007;**64**:175–81.
13. Ashcroft DM, Cooke J. Retrospective analysis of medication incidents reported using an on-line reporting system. *Pharm World Sci* 2006;**28**:359–65.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**327**:307–10.
15. Lawton R, Parker D. Barriers to incident reporting in a health-care system. *Qual Saf Health Care* 2002;**11**:15–8.
16. Evans SM, Berry JG, Smith BJ *et al.* Attitudes and barriers to incident reporting: a collaborative hospital study. *Qual Saf Health Care* 2006;**15**:39–43.

Accepted for publication 23 July 2009