# Lessons to learn from the reporting of adverse events (AEs) in randomised controlled trials: a systematic review of published reports in four high impact journals

**Rachel Phillips[1], Lorna Hazell[2], Odile Sauzet[3] and Victoria Cornelius[1]**

**[1]School of Public Health, Imperial College London, [2]Drug Safety Research Unit, and [3] Bielefeld School of Public Health, University Bielefeld.**

## INTRODUCTION

Reporting practices for adverse events (AEs) in randomised controlled trials (RCTs) are inadequate [(1,2)]. We undertook a systematic review to ascertain **best practice** for the collection, reporting and analysis of AEs in RCTs.

## METHODS

We included:

- reports of original RCTs;
- where the intervention was a pharmacological product;
- published in the Lancet, the BMJ, the NEJM and JAMA;
- between September 2015 and September 2016.

## RESULTS

| SECTION | DATA ITEM | N=184 | |
|---|---|---|---|
| **Collection** | | n | % |
| | Method of AE collection | | |
| | Passive collection | 114 | 62.0 |
| | Prompted collection | 53 | 28.8 |
| | *Active screening methods* | | |
| | Clinical examinations | 153 | 83.2 |
| | Laboratory tests | 146 | 79.4 |
| | Timing of prompted collection specified (n=53) | 48 | 90.6 |
| | Timing of active collection specified (n=166) | 95 | 57.2 |
| **Planned analysis** | | | |
| | Analysis for AEs specified in the methods section | 57 | 31.0 |
| | Population for AE analysis specified | 82 | 44.6 |
| | Planned interim analysis with stopping criteria for safety | 5 | 2.7 |
| **Selection for reporting** | | | |
| | Subset of AEs reported | 164 | 89.0 |
| | Criteria used[1] | | |
| | Frequency threshold | 48 | 26.1 |
| | Severity threshold | 17 | 9.2 |
| | Serious events | 42 | 22.8 |
| | Related to treatment | 15 | 8.2 |

[1]Full list of criteria use not presented. Papers typically used multiple criteria so groups presented are not mutually exclusive.

***Table 1: Summary of collection, assessment and analysis methods***

How AEs are **'selected'** for inclusion in the article was **inconsistent** and **unclear**:

- 89% reported a subset of all AEs recorded.
- 3% of studies did not specify what selection criteria was use.

AEs that cause patients to **withdraw** can be useful indicators of **severity** and **impact** to patients:

- 80% of studies reported numbers that withdrew from the trial.
- 35% of these reported whether the withdrawals were due to AEs.

| DATA ITEM | N=184 | |
|---|---|---|
| | n | % |
| **Withdrawals reported** | 146 | 79.4 |
| Withdrawals due to AEs reported (n=146) | 51 | 34.9 |
| Specific AEs causing withdrawals reported (n=51) | 12 | 23.5 |
| **Binary outcomes** | | |
| Binary AE outcomes summarised by arm | | |
| Not summarised | 6 | 3.3 |
| Number of people with an event | 154 | 83.7 |
| Number of events | 11 | 6.0 |
| Both | 12 | 6.5 |
| Unclear | 1 | 0.5 |
| Frequencies of AEs reported by arm | 173 | 94.1 |
| Percentages of AEs reported by arm | 160 | 87.0 |
| Between arm differences & 95% CI of AEs reported | 37 | 20.1 |
| Statistical significance tests between arms on AEs reported | 86 | 46.8 |
| **Continuous outcomes** | | |
| Continuous AEs outcomes dichotomised for summaries | 136 | 72.9 |
| If continuous outcomes were left as continuous what between arm analyses was performed? (n=38) | | |
| Differences in measures of central tendency estimated with 95% CI | 15 | 39.4 |
| Between arm hypothesis tests performed | 26 | 68.5 |
| Graphical presentations of AE outcomes | 22 | 12.0 |
| Summaries of severity rating of AEs reported | 76 | 41.3 |
| Number of serious AEs reported | 134 | 72.8 |
| Duration of any AEs reported | 9 | 4.9 |
| Time of occurrence of any AEs reported | 52 | 28.3 |
| Multiplicity of events accounted for by significance tests | 3 | 1.6 |

***Table 2: Summary of analysis performed and results presented***

Reporting numbers that experience at least one event and **ignoring repeated events** masks valuable information that may be **important to patients**:

- 84% of studies provided no information on the number of events occurring.

RCTs are typically **underpowered** to detect important differences in AEs between arms:

- 47% performed formal hypothesis tests for binary outcomes.

**Graphs** are an efficient method to **convey** and **interpret** large amounts of data and can make it easier to flag potential safety signals:

- 12% of studies used graphs to present AE data.

## CONCLUSION

This review has demonstrated that even in best practice the **collection**, **reporting** and **analysis** of AE data in clinical trials is sub-optimal.

Areas to improve include:

- Reducing information loss when analysing at patient level;
- Stopping the inappropriate practice of underpowered multiple hypothesis testing;
- Development of better statistical methods for AE analysis.

***REFERENCES**: 1. Ioannidis JPA et al. The Lancet 1998;* 352: *1752-1753. 2. Ioannidis JPA et al. JAMA 2001; 285: 437-443*

*Rachel Phillips is funded by the National Institute for Health Research (NIHR) Doctoral Research Fellowship programme to undertake this work. This is a summary of independent research and the views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.*