International Journal of Infection Control
www.ijic.info ISSN 1996-9783

ORIGINAL ARTICLE

What are these results telling me?

Elizabeth Anne Scicluna
Infection Control Unit, Mater Dei Hospital, Msida, Malta

doi: 10.3396/ijic.v8i4.034.12

Abstract

The scientific community must be able to read the published literature with a critical eye and look for the relevant information in the paper to see if they come up with the same conclusions as the authors. They should be aware of possible sources of bias and have enough information about the surveillance system used to collect the data. They should also be aware that meta-analysis is affected by publication bias and the reader should be able to determine how the authors selected the articles for their meta-analysis and if any conflict of interest present in the original papers have been reported. Several tools/checklists and guidelines were issued to ensure that the published data is accurate and reliable.

Key words

Infection control+methods; Journalisms, medical; Publication bias; Data interpretation, statistical

Corresponding Author
Elizabeth Anne Scicluna
Infection Control Unit, Mater Dei Hospital, Msida, Malta
Email: elizabeth.a.scicluna@gov.mt

Introduction

One of the things that an infection control practitioner should be able to do is 'Collate, analyse and communicate data relating to preventing and controlling infection for surveillance purposes'.[1] This include the knowledge to assess strengths and weaknesses of the data and use appropriate methods to analyse it and be able to determine whether the conclusions reached are really representing the populations from which this data are coming. However, many health care workers might not have the skills required to carefully examine and interpret statistical results and they assume that the results and interpretation given by the authors in the published literature are correctly reported.[2] In a review of 21 surveys that questions the frequency with which scientists fabricate and falsify data, 2% of scientists admitted to have fabricated, falsified or modified data or results and 34% reported other questionable research practices.[3] In the same review, when asking about the misconduct of colleagues, 14% reported that they know someone who falsify data and up to 72% reported research malpractice. The author concluded that since this is a sensitive subject the actual rates of falsification and fabrication of data might be even higher. The scientific community should be well aware of this and must be able to read the published literature with a critical eye and look for the relevant information in the paper to see if they come up with the same conclusions as the authors. Awareness of research misconduct has led to a number of ethical considerations to control the problem, such as through training of researchers, tougher scrutiny by the funding agencies and research institutions and finally by the editorial boards when the manuscript is sent out for publication.[4-6]

The aim of this paper is to help the reader to better interpret surveillance and research results that are frequently published either in scientific journals or as reports from the readers' own hospitals or organisations and also to highlight possible sources of errors in the data, such as bias and decisions taken from small number of observation.

Case definitions, Incidence and Prevalence

During surveillance, if one is monitoring trends but the case definition used over time changes it is impossible to make comparison. So standard case definitions need to be followed. In surveillance, consistency and timeliness is more important than high level of accuracy. In research studies, if one wishes to compare the results obtained with those from other studies, it is vital to use same case definitions.

Several agencies issue standardised definitions for different purposes. The International Classification of Diseases (ICD) is the international standard diagnostic classification for all general epidemiological and clinical use. ICD-10 is the latest version of these classification and standard case definitions.[7] The Centers for Disease Prevention and Control (CDC) has a set of case definitions for Health Care Associated Infections (HCAI).[8] Similar case definitions were issued by the ECDC and are used for both the longitudinal surveillance of specific HAI and also during the Point Prevalence Studies carried out throughout European hospitals.[9]

The number of new cases during a time period is the incidence, while prevalence is the number of existent cases at a point in time. So prevalence is affected by issues that increase or decrease the emergence of new cases and also by the duration of each case in that population. These factors may vary between populations or even in the same population in different time periods. For this reason the use of prevalence studies are not ideal to compare a situation between different populations. The use of prevalence studies is needed when one wants to determine the burden of a disease in a particular population to plan what facilities are needed.

One cannot compare absolute figures (number of cases) from different populations and it is tricky to try to compare absolute figures in the same population over time because you need to assume that all other possible predictor variables had remained equal. In such cases rates need to be calculated and so it is important to have the denominator data, which is information on the population at risk from which the cases have been identified.

Bias and Confounding

Bias is one of the main types of error that may occur in epidemiological studies. It refers to any error in

the design or conduct of a study, that results in a conclusion that is different from the truth. Bias arises because of problems with the study design. A biased study is one that does not give a true representation of the situation we want to describe or the association we want to analyse. It is particularly important that potential sources of bias are identified at the stage of study design because you cannot usually adjust or make allowance for bias at the analysis stage. Bias can ruin a study irretrievably.

There are different types of bias. The following are the main ones:

1. Selection bias

In descriptive studies (cross-sectional or cohort), selection bias occurs if the study population is not representative of the reference population, for example someone study hospitalised patients for convenience but intend to generalise the information on the general population.

The potential sources of selection bias in analytic studies depend on the type of study, but is mainly that the comparison groups are not comparable. In case-control studies, bias arises if cases are not representative of all cases within a defined population, for example one chooses to study less severe cases or cases that show up at a particular clinic only, or controls are not representative of the population which produced the cases. Selection bias in cohort studies may arise if the comparison groups (exposed and unexposed) are not truly comparable. This could arise because of poor choice of the unexposed group or due to differences in follow-up or case ascertainment between the two groups (example if more efforts are put in following the exposed to determine their outcome but not as much for the unexposed).

2. Observer bias

Observer bias arises in case-control or cross-sectional studies, when the accuracy of exposure data recorded by the investigator differs systematically between subjects in different outcome groups, example if the investigator keeps on asking and insisting on particular exposures in those that have the outcome but not the same for the controls, or in cohort or intervention studies, when the accuracy of outcome data recorded by the investigator differs systematically between subjects in different exposure groups, example when the investigator is in doubt is more likely to record a set of symptoms as a 'case' if s/he knows that this person is exposed.

3. Reporting bias

Reporting bias arises when subjects with a specific health outcome report previous exposures with a different degree of accuracy to those without the outcome, example a person that acquired a HCAI might give more accurate information on length of stay in hospital, infection control practices by the health care workers etc., because s/he might have already spent time thinking why this has happened and maybe paid more attention on what was going on around him/her after this episode unlike other patients, or when subjects who have experienced a specific exposure report subsequent health events with a different degree of accuracy to those who have not experienced the exposure, example if someone knows he is taking a new drug s/he will be more attentive for possible side effects.

Confounding is about alternative explanations. If in an Epidemiological study one wants to establish if there is an association between an exposure and an outcome, confounding will be the situation where this association is entirely or partially due to another exposure. The first question to ask about any change detected by a surveillance system is "is it real?". There are many opportunities for artefacts (that is factors other than a genuine change in disease frequency which affect the number of cases reported) to arise in surveillance data. Appropriate interpretation of the data requires a good knowledge of the reporting system/s used, its strengths and weaknesses, and sources of artefact.

1. Changes in completeness: Surveillance data do not have to be complete to be useful, but in order to interpret them, we need some idea of how complete they are, and be aware of changes that may influence them, example if we know that only 60% of doctors report notifiable diseases, and this is constant in time, than one can still monitor trends. The problem arise if this proportion is increased or decreased because just by looking

at the numbers reported one can erroneously conclude that the number of cases are increasing or decreasing respectively.

Completeness may be influenced by things as diverse as a change in the person who is responsible for reporting the data, a change in the reporting form used if it is made shorter or easier to fill in and so completeness may improve (and vice versa) or the introduction of a new test: if a new, more sensitive laboratory test is introduced, there may be a sudden apparent increase in the number of cases reported even if there is no real change in the rate of disease in the population.

We can sometimes assess the completeness of a source of data by comparing it with data on the same disease from another source for example by comparing the number of cases of a disease reported by clinicians to the number reported from laboratories. We would not expect the number of cases reported by each method to be the same, but the trends in the two sources should follow the same pattern. Completeness is particularly important for rare diseases, where a small change in completeness may cause a relatively large change in the numbers of cases reported.

2. Reporting delays can cause apparent "epidemics" in surveillance data. For example, there may be important differences in reporting delay between clinical and laboratory notifications. Laboratory tests may need time for organisms to grow before cases can be reported.

3. Accuracy: the accuracy of diagnoses will affect the interpretation of surveillance data. We can improve accuracy by having a well-organised system with clear procedures and case-definitions. We also need to know whether our data are representative of the true distribution of disease in the population. It may be helpful to use data from more than one source.

We often compare surveillance data with either earlier data from the same geographical region, to determine trends over time or contemporary data from other regions. However, we need to be sure that the data are comparable, as already mentioned above under case definition. The method of data collection, for example whether surveillance was passive or active, may also affect comparability. Before and after studies should either have a contemporary control (no intervention) group or there should be sufficient observations for analysis as an interrupted time series.[10,11] The minimum number of data points is three before and three after the intervention. A general recommendation is for at least 12 monthly data points before and 12 monthly points after the intervention, although more data points and longer study periods provides even stronger evidence because trends, seasonal effects and natural sporadic variability can be better identified.[12]

Sample Size and Power

The sample size we choose to study will have an effect on the results we obtain. Let's say that the intervention under study has little or no effect on the outcome of interest. The difference observed in a study is therefore likely to be non-significant. However, the width of the confidence interval for the effect measure (for example, the risk ratio) will depend on the sample size. If the sample is small, the confidence interval will be very wide, and so even though it will probably include the null value (a zero difference between the groups, or a risk ratio of 1), it will extend to include large values of the effect measure. In other words, the study will have failed to establish that the intervention has no appreciable effect.

Suppose now that the intervention does have an appreciable effect. A study that is too small will have low power; that is it will have little chance of giving a statistically significant difference. In other words there is little chance of being able to demonstrate that the intervention has an effect. Even if a significant difference is found, the confidence interval on the effect will still be very wide, so there will be uncertainty at the end of the study whether the effect of the intervention is small and unimportant, or very large and of major importance.

The conduct of trials that are too small has consequences extending beyond the results of the specific trial. There is considerable evidence that

studies showing large effects are more likely to be published than those showing little or no effect. Suppose a number of small trials of a specific intervention are conducted. Because of the large sampling error implied by small sample sizes, a few of these trials will produce estimates of the effect of the intervention that are much larger than the true effect. These trials are more likely to be published, and the result is that the findings in the literature are likely to overestimate considerably the true effects of interventions. This publication bias is much smaller for larger trials, because an adequate sample size means that such trials will give results that are much closer to the true effect and, in addition, a large trial showing little or no effect is more likely to be published than a small trial with a similar difference.

Conclusions based only on a review of published data should be interpreted cautiously, especially for observational studies. Studies with statistically significant results are more likely to be published than those finding no difference between the study groups and those with significant results are also more likely to lead to a greater number of publications in journals with a high citation impact factor.[13]

Meta-analysis

To try to make some sense from several, possibly small studies, with different results, meta-analysis combine all these studies, give weight to the sample size and try to come out with an effect of the association under study. However, publication bias may also affect meta-analysis. This 'file drawer problem' results in effect sizes that are biased, skewed or completely cut off, creating a serious base rate fallacy, in which the significance of the published studies is overestimated.[14] Funnel plots, which are scatter plots of sample size and effect sizes, are usually used to detect such biases; asymmetrical plots are interpreted to suggest that biases are present and should be interpreted cautiously.[15,16] Figure 1 shows a funnel plot on the right as would be expected if there was no publication bias while the one on the left shows a funnel plot with the 'file drawer problem' where only studies that obtain a positive effect are published. Large studies with minimal effect are more likely to be published and this can be seen in the funnel plot on the left. In smaller studies a larger effect than it is in reality can be observed due to larger sampling error and by chance. Again small studies with large effect are more often accepted for publication than same size studies with no effect.

Roseman *et al.* reported that conflicts of interests in the studies underlying the meta-analyses were rarely disclosed.[17] They reviewed 29 meta-analyses reports that together covered 509 randomized controlled trials (RCTs). Of these, 318 RCTs reported funding sources with 219 (69%) industry funded. One hundred and thirty-two of the 509 RCTs reported author conflict of interest disclosures, with 91 studies (69%) disclosing

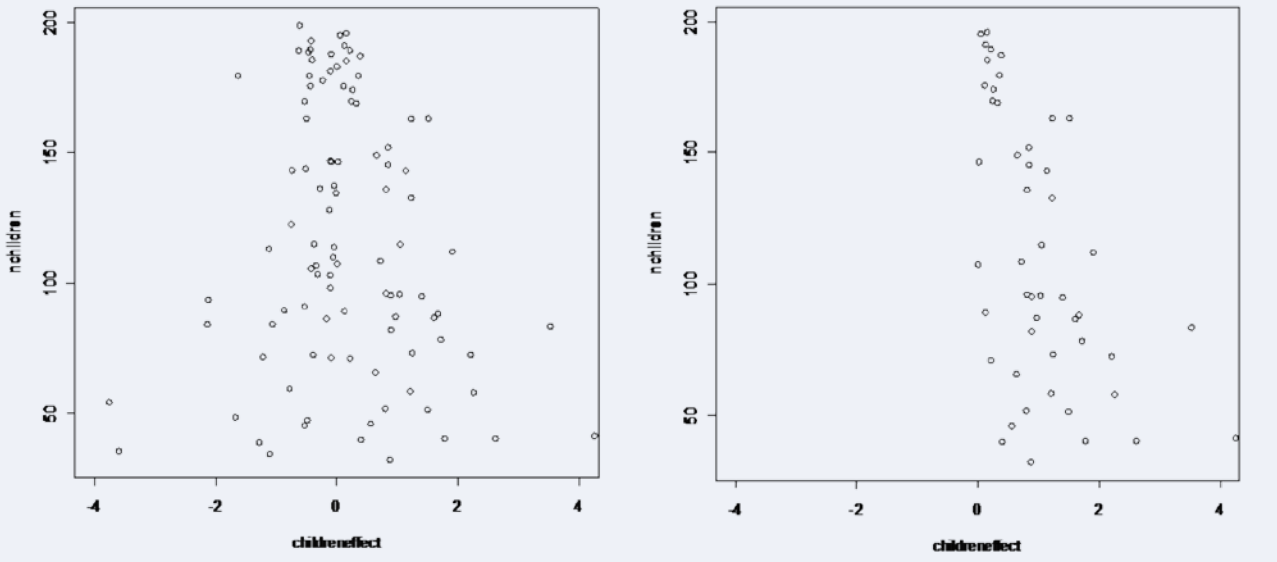

Figure 1. Funnel plots with and without the 'file drawer problem'
Source: http://en.wikipedia.org/wiki/Meta-analysis

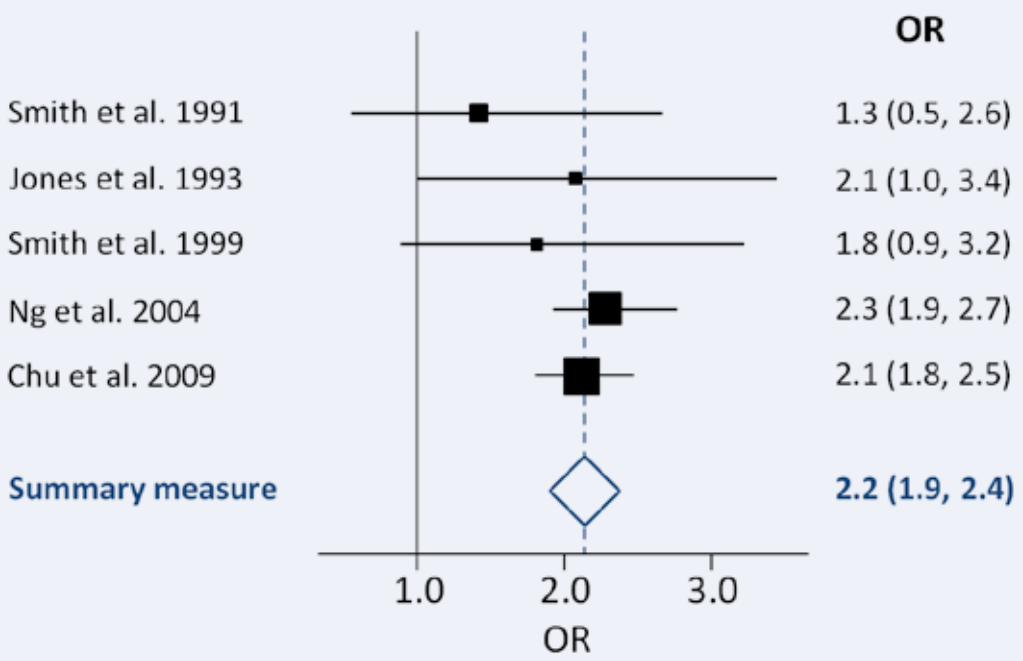

Figure 2. An example forest plot of five odds ratios (squares, proportional to weights used in meta-analysis), with the summary measure (centre line of diamond) and associated confidence intervals (lateral tips of diamond), and solid vertical line of no effect. Names of fictional studies are shown on the left, odds ratios and confidence intervals on the right
Source: http://en.wikipedia.org/wiki/Forest_plot

industry financial ties with one or more authors. However, very rarely was this information reflected in the meta-analyses. This was also reported in a Cochrane collection report.[18] Sismondo and Doucet presented evidence that pharmaceutical companies manipulate scientific literature, by controlling or shaping several crucial steps in the research, writing, and publication of scientific articles.[19] They also suggest that if medical journals want to ensure that the research they publish is ethically sound, they should not publish articles that are commercially sponsored.

Stroup and colleagues proposed a checklist which contains specifications for reporting of meta-analysis of observational studies in epidemiology.[20] These include background information, search strategy, methods, results, discussion, and conclusion. Use of this checklist should improve the usefulness of meta-analyses for authors, reviewers, editors, readers, and decision makers. Meta-analysis leads to a shift of emphasis from single studies to multiple studies. It emphasizes the practical importance of the effect size instead of the statistical significance of individual studies. The results of a meta-analysis are often shown in a forest plot (Figure 2). Results from studies are combined using different approaches. A common approach in health care research is to compute a weighted mean. Larger studies and studies with less random variation are given greater weight than smaller studies.

Conclusions

Several tools/checklists and guidelines were issued to ensure that the published data is accurate and reliable. Readers should be familiar with these to make better judgment of what they are reading. These include the tool for assessing the quality of published observational research that was developed by Sanderson *et al.* after reviewing the literature.[21] The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Initiative developed recommendations on what should be included in an accurate and complete report of an observational study.[22] Similarly ORION (Outbreak Reports and Intervention studies of Nosocomial infection) guidelines were produced to raise the standards of research and publication in hospital epidemiology, to facilitate synthesis of evidence and promote transparency of reporting,

to enable readers to relate studies to their own experience and assess the degree to which results can be generalised to other settings.[23] More recently, a revised CONSORT (Consolidated Standards of Reporting Trials) statement was published.[24] These are guidelines on the design and conduct of Randomised Controlled Trails.

References

1. Burnett E. Outcome competences for practitioners in infection prevention and control. *Journal of Infection Prevention* 2011; **12(2):** 67-90. http://dx.doi.org/10.1177/1757177410395797
2. Ferrill MJ, Brown DA, Kyle JA. Clinical versus statistical significance: interpreting P values and confidence intervals related to measures of association to guide decision making. *J Pharm Pract* 2010; **23(4):** 344-351. http://dx.doi.org/10.1177/0897190009358774
3. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One* 2009; **4(5):** e5738. http://dx.doi.org/10.1371/journal.pone.0005738
4. Coultas D. Ethical considerations in the interpretation and communication of clinical trial results. *Proc Am Thorac Soc* 2007; **4(2):** 194-198. http://dx.doi.org/10.1513/pats.200701-007GC
5. Kansu E, Ruacan S. Research ethics and scientific misconduct in biomedical research. *Acta Neurochir Suppl* 2002; **83:** 11-15.
6. Cehreli M, Cehreli Z, Stamm T, Meyer U, Wiesmann HP. Trick or treat? *Head Face Med* 2007; **3:** 22. http://dx.doi.org/10.1186/1746-160X-3-22
7. World Health Organisation. International Statistical Classification of Diseases and Related Health Problems 10th Revision. http://apps.who.int/classifications/icd10/browse/2010/en (accessed 7th May 2012).
8. Center for Disease prevention and control. CDC/NHSN Surveillance Definition of Healthcare-Associated Infection and Criteria for Specific Types of Infections in the Acute Care Setting. http://www.cdc.gov/nhsn/PDFs/pscManual/17pscNosInfDef_current.pdf (accessed 7th May 2012).
9. European Centre for Disease Prevention and Control. ECDC point prevalence survey of healthcare-associated infections and antimicrobial use in acute care hospitals. http://www.ecdc.europa.eu/en/activities/surveillance/HAI/about_HAI-Net/Pages/PPS.aspx (accessed 7th May 2012).
10. Ramsay C, Brown E, Hartman G, Davey P. Room for improvement: a systematic review of the quality of evaluations of interventions to improve hospital antibiotic prescribing. *Journal of Antimicrobial Chemotherapy* 2003; **52:** 764-771. http://dx.doi.org/10.1093/jac/dkg460
11. Davey P, Brown E, Hartman G, *et al*. Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database Syst Rev* 2005; **19(4):** CD003543. http://dx.doi.org/10.1002/14651858.CD003543.pub2
12. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharmac Therapeut* 2002; **27:** 299-309. http://dx.doi.org/10.1046/j.1365-2710.2002.00430.x
13. Easterbrook PJ, Berlin J, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991; **337:** 867–872. http://dx.doi.org/10.1016/0140-6736(91)90201-Y
14. Wikipedia. Meta-analysis. http://en.wikipedia.org/wiki/Meta-analysis (accessed 7th May 2012).
15. Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *J Clin Epidemiol* 2000; **53:** 477–484. http://dx.doi.org/10.1016/S0895-4356(99)00204-8
16. Biljana M, Jelena M, Branislav J, Milorad R. Bias in meta-analysis and funnel plot asymmetry. *Stud Health Technol Inform* 1999; **68:** 323-328.
17. Roseman M, Milette K, Bero LA, *et al*. Reporting of conflicts of interest in meta-analyses of trials of pharmacological treatments. *JAMA* 2011; **305(10):** 1008-1017. http://dx.doi.org/10.1001/jama.2011.257
18. The Cochrane Collaboration. How well do meta-analyses disclose conflicts of interests in underlying research studies. http://www.cochrane.org/news/blog/how-well-do-meta-analyses-disclose-conflicts-interests-underlying-research-studies (accessed 7th May 2012).
19. Sismondo S, Doucet M. Publication ethics and the ghost management of medical publication. *Bioethics* 2010; **24(6):** 273-283. http://dx.doi.org/10.1111/j.1467-8519.2008.01702.x
20. Stroup DF, Berlin JA, Morton SC, *et al*. Meta-analysis of Observational Studies in Epidemiology - A Proposal for Reporting. *JAMA* 2000, **283(15):** 2008-2012. http://dx.doi.org/10.1001/jama.283.15.2008
21. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007; **36:** 666-676. http://dx.doi.org/10.1093/ije/dym018
22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP for the STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology* 2008; **61:** 344-349. http://dx.doi.org/10.1016/j.jclinepi.2007.11.008
23. Stone SP, Cooper BS, Kibbler CC, *et al*. The ORION statement: guidelines for transparent reporting of outbreak reports and intervention studies of nosocomial infection. *Journal of Antimicrobial Chemotherapy* 2007; **59:** 833–840. http://dx.doi.org/10.1093/jac/dkm055
24. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Annals of Internal Medicine* 2010; **152(11):** 726-732.