GLOSSARY

Bias

Miguel Delgado-Rodríguez, Javier Llorca

J Epidemiol Community Health 2004;**58**:635–641. doi: 10.1136/jech.2003.008466

The concept of bias is the lack of internal validity or incorrect assessment of the association between an exposure and an effect in the target population in which the statistic estimated has an expectation that does not equal the true value. Biases can be classified by the research stage in which they occur or by the direction of change in a estimate. The most important biases are those produced in the definition and selection of the study population, data collection, and the association between different determinants of an effect in the population. A definition of the most common biases occurring in these stages is given.

See end of article for authors' affiliations

Correspondence to: Professor M Delgado-Rodríguez, Division of Preventive Medicine and Public Health, Building B-3, University of Jaen, 23071-Jaén, Spain; mdelgado@ujaen.es

Accepted for publication 15 November 2003

The concept of bias is the lack of internal validity or incorrect assessment of the association between an exposure and an effect in the target population. In contrast, external validity conveys the meaning of generalisation of the results observed in one population to others. There is not external validity without internal validity, but the presence of the second does not guarantee the first. Bias should be distinguished from random error or lack of precision. Sometimes, the term bias is also used to refer to the mechanism that produces lack of internal validity.[1]

Biases can be classified by the direction of the change they produce in a parameter (for example, the odds ratio (OR)). Toward the null bias or negative bias yields estimates closer to the null value (for example, lower and closer OR to 1), whereas away from the null bias produces the opposite, higher estimates than the true ones. An exaggeration of these biases can induce a switch-over bias, or change of the direction of association (for example, a true OR >1 becomes <1).[2]

There are several classifications of bias. Sackett[3] and Choi[4] classified biases according to the stages of research that can occur: reading up on the field, specification and selection the study sample, execution of the experimental manoeuvre, measurement of exposures/outcomes, data analysis, results interpretation and publication. Maclure and Schneeweiss,[5] applying the causal diagram theory, offered an interesting explanation of the main sources of bias. Kleinbaum *et al*,[2] based on Olli S Miettinen's ideas, classified biases in three main groups: selection bias, information bias, and confounding. Steineck and Ahlbom,[6] based on the Miettinen's concept of study base, considered in this order, confounding, misclassification (similar to information bias), misrepresentation (which has a narrower meaning than selection bias), and analysis deviation. Steineck and Ahlbom keep confounding apart from biases in the statistical analysis as it typically occurs when the actual study base differs from the ''ideal'' study base, in which there is no association between different determinants of an effect. The same idea can be found in Maclure and Schneeweiss.[5]

In this glossary definitions of the most common biases (we have not been exhaustive in defining all the existing biases) are given within the simple classification by Kleinbaum *et al*.[2] We have added a point for biases produced in a trial in the execution of the intervention. Biases in data interpretation, writing, and citing will not be discussed (see for a description of them by Sackett[3] and Choi[4]). Biases in data analysis are very numerous, but they are easily solved using appropriate procedures; they are not commented on in this glossary, unless a particular bias has an additional influence. This occurs with post hoc analysis that may lead to a publication bias when significant results are more frequently reported. Confounding bias is kept apart from biases in data analysis (according to the ideas of Steineck and Ahlbom[6] and Maclure and Schneeweiss[5]).

Table 1 gives an alphabetical list of biases. The type of bias and the design affected is also given.

1 SELECTION BIAS

The error introduced when the study population does not represent the target population.[7 8] Selection bias can be controlled when the variables influencing selection are measured on all study subjects and either (a) they are antecedents of both exposure and outcome or (b) the joint distribution of these variables (plus exposure and outcome) is known in the whole target population, or (c) the selection probabilities for each level of these variables are known.[9] It can be introduced at any stage of a research study[7]: design (bad definition of the eligible population, lack of accuracy of sampling frame, uneven diagnostic procedures in the target population) and implementation.

1.1 Inappropriate definition of the eligible population

In any kind of design *ascertainment bias* can occur. It is produced when the kind of patients gathered does not represent the cases originated in the population[4] (see Pollock *et al*[10] for an illustration). It may be produced, among many possibilities, by healthcare access bias, length-biased sampling, Neyman bias, competing risks, or survivor treatment selection bias. In studies on evaluation of a diagnostic test the spectrum bias is a kind of ascertainment bias. The definitions of these biases in alphabetical order are the following:

Table 1 Alphabetical list of biases, indicating their type and the design where they can occur

Specific name of bias	Group of bias	Subgroup of bias (next level to specific name)	Type of design affected
Allocation of intervention bias	Execution of an intervention		Trial
Apprehension bias	Information bias	Observer bias	All studies
Ascertainment bias	Selection bias	Inappropriate definition of the eligible population	Observational study
Berkson's bias	Selection bias	Inappropriate definition of the eligible population	Hospital based case-control study
Centripetal bias	Selection bias	Healthcare access bias	Observational study
Citation bias	Selection bias	Lack of accuracy of sampling frame	Systematic review/meta-analysis
Competing risks	Selection bias	Ascertainment bias	All studies
Compliance bias	Execution of an intervention		Trial
Confounding by group	Confounding		Ecological study
Confounding by indication	Confounding		Case-control study, cohort study
Contamination bias	Execution of an intervention		Trial, mainly community trials
Detection bias	Selection bias	Uneven diagnostic procedures in the target population	Case-control study
Detection bias	Information bias	Misclassification bias	Cohort study
Diagnostic/treatment access bias	Selection bias	Healthcare access bias	Observational study
Diagnostic suspicion bias	Selection bias	Detection bias	Case-control study
Diagnostic suspicion bias	Information bias	Detection bias	Cohort study
Differential maturing			Trial
Differential misclassification bias	Information bias	Misclassification bias	All studies
Dissemination bias	Selection bias	Lack of accuracy of sampling frame	Systematic review/meta-analysis
Ecological fallacy	Information bias		Ecological study
Exclusion bias	Selection bias	Inappropriate definition of the eligible population	Case-control study
Exposure suspicion bias	Information bias	Recall bias	Case-control study
Family aggregation bias	Information bias	Reporting bias	Observational study
Friend control bias	Selection bias	Inappropriate definition of the eligible population	Case-control study
Hawthorne effect	Information bias		Trial
Healthcare access bias	Selection bias	Ascertainment bias	Observational study
Healthy volunteer bias	Selection bias	Non-response bias	Observational study
Healthy worker effect	Selection bias	Inappropriate definition of the eligible population	Cohort study (mainly retrospective)
Incidence-prevalence bias (synonym of Neyman bias)			
Inclusion bias	Selection bias	Inappropriate definition of the eligible population	Hospital based case-control study
Lack of intention to treat analysis			Randomised trial
Language bias	Selection bias	Inappropriate definition of the eligible population	Systematic review/meta-analysis
Lead-time bias	Information bias		Screening study
Length biased sampling	Selection bias	Ascertainment bias	Cross sectional study, screening
Losses/withdrawals to follow up	Selection bias	During study implementation	Cohort study, trial
Mimicry bias	Selection bias	Detection bias	Case-control study
Mimicry bias	Information bias	Detection bias	Cohort study
Misclassification bias	Information bias		All studies
Missing information in multivariable analysis	Selection bias	During study implementation	All studies (mainly retrospective)
Mode for mean bias	Information bias	Reporting bias	All studies
Neyman bias	Selection bias	Ascertainment bias	Cross sectional study, case-control study with prevalent cases
Non-differential misclassification bias	Information bias	Misclassification bias	All studies
Non-random sampling bias	Selection bias	Lack of accuracy of sampling frame	Observational study
Non-response bias	Selection bias	During study implementation	Observational study
Obsequiousness bias	Information bias	Reporting bias	All studies
Observer expectation bias	Information bias	Observer bias	All studies
Observer/interviewer bias	Information bias	Misclassification bias	All studies
Overmatching	Selection bias	Inappropriate definition of the eligible population	Case-control study
Participant expectation bias	Information bias	Recall bias	Trial
Popularity bias	Selection bias	Healthcare access bias	Observational study
Post hoc analysis	Selection bias	Publication bias	Systematic review/meta-analysis
Protopathic bias	Information bias		Observational study
Publication bias	Selection bias	Lack of accuracy of sampling frame	Systematic review/meta-analysis
Purity diagnostic bias	Selection bias	Spectrum bias	Validity of diagnostic tests
Recall bias	Information bias	Misclassification bias	All studies
Referral filter bias	Selection bias	Healthcare access bias	Observational study
Regression dilution bias	Information bias	Regression to the mean	Cohort study, trial
Regression to the mean	Information bias		Cohort study, trial
Relative control bias	Selection bias	Inappropriate definition of the eligible population	Case-control study
Reporting bias	Information bias	Misclassification bias	All studies
Rumination bias	Information bias	Recall bias	Case-control study, retrospective cohort study

Table 1 Continued

Specific name of bias	Group of bias	Subgroup of bias (next level to specific name)	Type of design affected
Selective survival bias (synonym of Neyman bias)			
Sick quitter bias	Information bias	Protopathic bias	Observational study
Spectrum bias	Selection bias	Ascertainment bias	Validity of diagnostic tests (mainly case-control study)
Survivor treatment selection bias	Selection bias	Ascertainment bias	Cohort study (mainly retrospective)
Susceptibility bias (synonym of confounding)			
Telephone random sampling bias	Selection bias	Non-random sampling bias	Observational study
Temporal ambiguity	Information bias		Cross sectional study, ecological study
Unacceptable disease/exposure	Information bias	Reporting bias	Observational study
Underreporting bias	Information bias	Reporting bias	Observational study
Unmasking—detection signal—bias	Selection bias	Detection bias	Case-control study
Verification bias (synonym of work up bias)			
Will Rogers phenomenon	Information bias		Prognostic (mainly cohort) study
Work up bias	Information bias		Validity of diagnostic test (retrospective study)

- *Competing risks:* when two or more outputs are mutually exclusive, any of them competes with each other in the same subject. It is more frequent when dealing with causes of death: as any person only dies once, the risk for a specific cause of death can be affected by an earlier one. For example, early death by AIDS can produce a decrease in liver failure mortality in parenteral drug users. A proper analysis of this question should take into account the competing causes of death; for instance, estimating the probability of death by a specific cause of death if any other risk of death is removed (the so called net probability of death).[11 12]
- *Healthcare access bias:* when the patients admitted to an institution do not represent the cases originated in the community. This may be due: to the own institution if admission is determined by the interest of health personnel on certain kind of cases (*popularity bias*),[3] to the patients if they are attracted by the prestige of certain clinicians (*centripetal bias*),[3] to the healthcare organisation if it is organised in increasing levels of complexity (primary, secondary, and tertiary care) and ''difficult'' cases are referred to tertiary care (*referral filter bias*),[3] to a web of causes if patients by cultural, geographical, or economic reasons show a differential degree of access to an institution (*diagnostic/treatment access bias*).[3]
- *Length-bias sampling:* cases with diseases with long duration are more easily included in surveys. This series may not represent the cases originated in the target population.[13] These cases usually have a better prognosis.
- *Neyman bias:* (synonyms: *incidence-prevalence bias*, *selective survival bias*) when a series of survivors is selected, if the exposure is related to prognostic factors, or the exposure itself is a prognostic determinant, the sample of cases offers a distorted frequency of the exposure.[14] This bias can occur in both cross sectional and (prevalent) case-control studies. Lets suppose that a case-control study is carried out to study the relation between tobacco smoking and acute myocardial infarction (AMI), being cases interviewed one week after the coronary attack. If smoker patients with AMI die more frequently, the leaving cases will show lower frequency of smoking, undervaluing the association between smoking and AMI. It has been shown that the bias occurs only if the risk factor influences mortality from the disease being studied.[15]
- *Spectrum bias:* in the assessment of validity of a diagnostic test this bias is produced when researchers included only ''clear'' or ''definite'' cases, not representing the whole spectrum of disease presentation, and/or ''clear'' or healthy controls subjects, not representing the conditions in which a differential diagnosis should be carried out. Sensitivity and specificity of a diagnostic test are increased.[16] A particular case is the *purity diagnostic bias*, when selecting cases of a certain disease those with other comorbidities are excluded and the final sample does not represent the cases originated.[3]
- *Survivor treatment selection bias:* in observational studies patients who live longer have more probability to receive a certain treatment. A retrospective analysis can therefore yield a positive association between that treatment and survival.[17]

In occupational studies a bad definition, albeit unavoidable, of eligible population frequently occurs. It is produced by the *healthy worker effect*: the lower mortality observed in the employed population when compared with the general population. Furthermore, those who remain employed tend to be healthier than those who leave employment.[18]

Inadequate definition of the eligible population can happen frequently in case-control studies, where several specific biases have been described regarding the selection of the reference population (controls). The most common are (in alphabetical order):

- *Berkson's bias:* first described by Berkson in 1946 for case-control studies.[19] It is produced when the probability of hospitalisation of cases and controls differ, and it is also influenced by the exposure. See Feinstein *et al*[20] and Flanders *et al*[21] for a discussion and solution.
- *Exclusion bias:* when controls with conditions related to the exposure are excluded, whereas cases with these diseases as comorbidities are kept in the study. This was the explanation given for the association between reserpine and breast cancer: controls with cardiovascular disease (a common comorbidity and related to the use of reserpine) were excluded but this criterion was not applied to cases, thus yielding a spurious association between reserpine and breast cancer.[22]
- *Friend control bias:* It was assumed that the correlation in exposure status between cases and their friend controls lead to biased estimates of the association between exposure and outcome. In a matched study, with a matched analysis, there is no bias if the exposure induced risks of disease are constant over time and there are not gregarious subjects, individuals elected by more than one case.[23]

- *Inclusion bias:* produced in hospital based case-control studies when one or more conditions of controls are related with the exposure. The frequency of exposure is higher than expected in the reference group, producing a toward the null bias.[1]
- *Matching:* It is well known that matching, either individual or frequency matching, introduces a selection bias, which is controlled for by appropriate statistical analysis: matched analysis in studies with individual matching and adjusting for the variables used to match in frequency matching. Overmatching is produced when researchers match by a non-confounding variable (associated to the exposure but not to the disease) and can underestimate an association.[1]
- *Relative control bias:* It was assumed that the correlation in exposure status between cases and their relative controls yield biased estimates of the association between exposure and outcome. In a matched study, with a matched analysis, there is no bias if the exposure induced risks of disease are constant over time.[24]

In systematic reviews and meta-analyses *language bias* is a kind of inappropriate definition of the eligible population (the reports studying the relevant topic). In systematic reviews and meta-analysis it has been common to exclude reports in other languages than English. Egger *et al*[25] showed that there was a trend to publish in English compared with German when the results achieved statistical significance; later on, the same group found that language bias has in general little effect on summary effect estimates.[26]

1.2 Lack of accuracy of sampling frame

The most common bias in this group is *non-random sampling bias*: obviously, this selection procedure can yield a non-representative sample in which a parameter estimate differs from the existing at the target population.[4] A particular case of this bias is *telephone random sampling bias*: it excludes some households from the sample, thus producing a coverage bias. In the US it has been shown that the differences between participants and non-participants are generally not large,[27] but the situation can be very different in less developed countries.

In systematic reviews and meta-analyses focused only in published reports the most important selection biases are publication bias and other biases influencing the identification of relevant studies (citation bias and dissemination bias):

- *Citation bias:* articles more frequently cited are more easily found and included in systematic reviews and meta-analysis. Citation is closely related to the impact factor of the publishing journal.[28] In certain fields, citation has been related to statistical significance.[29]
- *Dissemination bias:* the biases associated to the whole publication process, from biases in the retrieval of information (including language bias) to the way the results are reported.[30]
- *Post hoc analysis:* the fishing expeditions with data dredging originates post hoc questions and subgroup analysis with misleading results.[4] Given that the reports based on post hoc analysis are frequently reported when significant results are observed, this bias is relevant for meta-analysis of published studies as a form of publication (selection) bias.
- *Publication bias:* regarding an association that is produced when the published reports do not represent the studies carried out on that association. Several factors have been found to influence publication, the most important being statistical significance, size of the study, funding, prestige, type of design, and study quality.[31]

1.3 Uneven diagnostic procedures in the target population

In case-control studies, if exposure influences the diagnosis of the disease, *detection bias* occurs. Particular types of this bias are: exposure can be taken as another diagnostic criterion (*diagnostic suspicion bias*).[3] Exposure can trigger the search for the disease; for instance, benign anal lesions increases the diagnosis of anal cancer.[32] Exposure may produce a symptom/sign that favours diagnosis (*unmasking-detection signal-bias*)[3] or a benign condition close clinically to the disease (*mimicry bias*).[3] In other designs (such as cohort studies) detection bias is an information bias.

1.4 During study implementation

The three most common biases at this stage are losses to follow up, missing information in multivariable analysis, and non-response bias:

- *Losses/withdrawals to follow up:* in both cohort and experimental studies when losses/withdrawals are uneven in both the exposure and outcome categories, the validity of the statistical results may be affected.[33]
- *Missing information in multivariable analysis:* multivariable analysis selects records with complete information on the variables included in the model. If participants with complete information do not represent the target population, it can introduce a selection bias.[7] This bias is relevant in studies, mainly retrospective, using data from the clinical chart, in which patients with more complete data have more severe diseases or stay longer at hospital, or both.
- *Non-response bias:* when participants differ from non-participants, for example, Melton *et al*.[34] The *healthy volunteer effect* is a particular case: when the participants are healthier than the general population.[35] This is particularly relevant when a diagnostic manoeuvre, such as a screening test, is evaluated in the general population, producing an away from the null bias; thus the benefit of the intervention is spuriously increased.

2 INFORMATION BIAS

Information bias occurs during data collection. The three main types of information bias are misclassification bias, ecological fallacy, and regression to the mean. Other information biases are also described.

2.1 Misclassification bias

It is originated when sensitivity and/or specificity of the procedure to detect exposure and/or effect is not perfect, that is, exposed/diseased subjects can be classified as non-exposed/non-diseased and vice versa.[36] Given that perfect tools to gather data are very uncommon most studies must assume a certain degree of misclassification. Random error also can produce it.[37] This implies that random errors in data entry/capture, missing data, end digit preference (rounding to 5 or 0), frequently unavoidable, also introduce misclassification. There are two major types of misclassification bias:

- *Differential misclassification bias:* when misclassification is different in the groups to be compared; for example, in a case-control study the recalled exposure is not the same for cases and controls. The estimate is biased in either direction, toward the null or away from the null.[36]
- *Non-differential misclassification bias:* when the misclassification is the same across the groups to be compared, for

example, exposure is equally misclassified in cases and controls. For binary variables the estimate is biased toward the null value[36]; however, for variables with more than two categories (polytomous) this rule may not hold and an away from the null bias can be obtained.[38]

The most common biases producing misclassification are the following

- *Detection bias* in studies with follow up (cohorts, clinical trials) is an information bias.
- *Observer/interviewer bias:* the knowledge of the hypothesis, the disease status, or the exposure status (including the intervention received) can influence data recording (*observer expectation bias*).[3] The means by which interviewers can introduce error into a questionnaire include administering the interview or helping the respondents in different ways (even with gestures), putting emphases in different questions, and so on.[39] A particular situation is when the measure of an exposure influences its value (for example, blood pressure) (*apprehension bias*).[3]
- *Recall bias:* if the presence of disease influences the perception of its causes (*rumination bias*)[3] or the search for exposure to the putative cause (*exposure suspicion bias*),[3] or in a trial if the patient knows what they receive may influence their answers (*participant expectation bias*).[3] This bias is more common in case-control studies, in which participants know their diseases, although it can occur in cohort studies (for example, workers who known their exposure to hazardous substances may show a trend to report more the effects related to them), and trials without participants' blinding.[40]
- *Reporting bias:* participants can ''collaborate'' with researchers and give answers in the direction they perceive are of interest (*obsequiousness bias*),[3] or the existence of a case triggers family information (*family aggregation bias*), see Khoury *et al*[41] for an example. Measures or sensitive questions that embarrass or hurt can be refused (*unacceptable disease/exposure*).[3] *Underreporting bias* is common with socially undesirable behaviours, such as alcohol consumption.[42] The *mode for mean bias* occurs when frequency-quantity questionnaires are used to assess consumption of alcohol and foods, subjects tend to report modal rather than average behaviour, hence with data skewed towards zero, the average intakes are underestimated, leading to overestimation of the gradient with risk.[43]

The last three biases can be reduced using blinding, a procedure by which subjects ignore some important aspects of a research to avoid differential misclassification bias. In trials, blinding means that participants do not know the intervention they receive (participants blinding) and/or observers do not know the intervention received by participants (observer blinding), and/or data analysts do not know the labels of the groups to be compared.[44] In observational research, observer blinding can rarely be applied with the same goal of trials, but more frequently observers and participants are blind to the main hypotheses (and related questions) of a research.

2.2 Ecological fallacy

It is a bias produced when analyses realised in an ecological (group level) analysis are used to make inferences at the individual level. For instance, if exposure and disease are measured at the group level (for example, exposure prevalence and disease risk in each country), exposure-disease relations can be biased from those obtained at the individual level (for example, exposure status and disease status in each subject). Ecological fallacy can be produced by within group (individual level) biases, such as confounding, selection bias, or misclassification, and by confounding by group or effect modification by group.[45] Effect modification by group on an additive scale is produced when the rate difference for the exposure effect changes across communities[45]; for example, lets suppose three groups with exposure prevalences of 35%, 50%, and 65%, a similar rate of disease in the non-exposed of 100/100 000, and rates of disease in the exposed of 286, 200, and 154/100 000. In this example, the rate difference in each group is positive (the exposure increases the risk of the disease), although an ecological analysis do find a negative relation.

2.3 Regression to the mean

It is the phenomenon that a variable that shows an extreme value on its first assessment will tend to be closer to the centre of its distribution on a later measurement.[46] This bias is relevant when the efficacy of a treatment to reduce high levels of a variable (for example, cholesterol) is assessed, when researchers are interested in the relation between the initial value of a variable and the change in that measurement over time, or when two methods of measurement are compared.[47] The two usual ways of neutralising this bias are with the existence of an appropriate reference group and a selection based on more than one measurement.

The *regression dilution bias* is related to the regression to the mean. This bias is produced in longitudinal studies analysing baseline determinations of a continuous variable (such as diastolic blood pressure (DBP)) to an outcome (for example, stroke). Baseline DBP measurements randomly fluctuates among individuals because of two reasons: variations in the measurement process and temporary deviations at the baseline determination from the usual DBP level. This underestimates the real association between exposure and outcome because extreme categories include more people than they should, that is, the bottom category of baseline DBP has more people whose DBP level is somewhat lower than their usual DBP, whereas the top category of DBP includes more subjects with higher baseline DBP than their usual DBP level.[48]

2.4 Other information biases

- *Hawthorne effect:* described in the 1920s in the Hawthorne plant of the Western Electric Company (Chicago, IL). It is an increase in productivity—or other outcome under study—in participants who are aware of being observed.[49] For example, laboratory physicians increase their agreement rate after knowing that they participate in a research on reliability of diagnostic tests.[50]
- *Lead time bias:* the added time of illness produced by the diagnosis of a condition during its latency period. This bias is relevant in the evaluation of the efficacy of screening, in which the cases detected in the screened group has a longer duration of disease than those diagnosed in the non-screened one.[51]
- *Protopathic bias:* when a exposure is influenced by early (subclinical) stages of disease. For instance, preclinical pancreatic cancer can produce diabetes mellitus, and thus an association between diabetes and cancer can occur.[52] It is also produced when a pharmaceutical agent is prescribed for an early manifestation of a disease that has not been yet diagnosed.[53] The *sick quitter bias* is related to protopathic bias: people with risky behaviours (such as heavy alcohol consumption) quit their habit as a consequence of disease[54]; studies analysing current behaviour as a risk factor will labelled them as non-exposed, thus underestimating the true association.

- *Temporal ambiguity:* when it cannot be established that exposure precedes effect. It is common in cross sectional and ecological studies.[1]
- *Will Rogers phenomenon:* named in honour of the philosopher Will Rogers by Feinstein *et al*.[55] The improvement in diagnostic tests refines disease staging in diseases such as cancer. This produces a stage migration from early to more advances stages and an apparent higher survival. This bias is relevant when comparing cancer survival rates across time or even among centres with different diagnostic capabilities (for example, tertiary compared with primary care hospitals).
- *Work up bias (verification bias)*: in the assessment of validity of a diagnostic test, it is produced when the execution of the gold standard is influenced by the results of the assessed test, typically the reference test is less frequently performed when the test result is negative.[16 56] This bias is aggravated when the clinical characteristics of a disease influence in the test results.[57]

3 CONFOUNDING

It occurs when a variable is a risk factor for an effect among non-exposed persons and is associated with the exposure of interest in the population from which the effect derives, without being affected by the exposure or the disease (in particular, without being an intermediate step in the causal pathway between the exposure and the effect).[1] The counterfactual approach is the current procedure to explain confounding adequately,[58] and causal diagrams help to identify it.[59] Confounding can occur in every epidemiological study. *Susceptibility bias* is a synonym: when people who are particularly susceptible to development of a outcome are also prone to be exposed; for example, women with threatened abortion have a high probability of delivering a malformed fetus but also have a high probability of receiving hormone treatment. This can yield a spurious association between hormones and congenital malformations.[53]

Confounding can be neutralised at the design stage of a research (for example, by matching or randomisation) and/or at the analysis, given that the confounders have been measured properly. Misclassification of confounders hinders their control in analysis. Non-differential misclassification of a binary confounder reduces the ability of analysis to control for the confounder,[60] whereas this in a polytomous confounder can produce estimates biased in either direction.[61] Particular types of confounding bias are the following:

- *Confounding by group:* it is produced in an ecological study, when the exposure prevalence of each community (group) is correlated with the disease risk in non-exposed of the same community. It can be a mechanism for producing ecological fallacy.[45] For example, lets suppose three communities (A, B, C) with prevalence exposures of 10%, 20%, and 30%, rates of disease in non-exposed of 2%, 3%, and 4%, and rates of disease in the exposed of 2%, 3%, 4%, respectively. There is no association between the exposure and the disease as the three relative risks are one, although an ecological analysis, regressing the rate of disease on the prevalence of exposure, does find a positive association.
- *Confounding by indication:* this is produced when an intervention (treatment) is indicated by a perceived high risk, poor prognosis, or simply some symptoms. Here the confounder is the indication, as it is related to the intervention and is a risk indicator for the disease.[62] For example, in the study of the association between cimetidine and gastric cancer, the indication peptic ulcer is considered the potential confounder.[63] This kind of bias occurs in observational studies (mainly retrospective) analysing interventions. Sometimes confounding by indication is mistaken for protopathic bias.

4 SPECIFIC BIASES IN TRIALS

- *Allocation of intervention bias:* when intervention is differentially assigned to the population. It is more common in non-randomised trials. In randomised trials it is recommended concealment of the allocation sequence of intervention.[64] If the sequence is known in advance may produce selection bias. It has been shown that trials in which concealment was unclear or inadequate, compared with trials with adequate concealment, report larger estimates of treatment effects.[65]
- *Compliance bias:* in trials requiring adherence to intervention, the degree of adherence (compliance) influences efficacy assessment of the intervention. For example, when high risk patients quit exercise programmes.[3]
- *Contamination bias:* when intervention-like activities find their way into the control group. It biases the estimate of the intervention effect toward the null hypothesis.[66] It occurs more frequently in community intervention trials because of the relationships among members of different communities and interference by mass media, health professionals, etc.
- *Differential maturing:* in group randomised trials differential maturing reflects uneven secular trends among the groups in the trial favouring one condition or another.[66]
- *Lack of intention to treat analysis:* in randomised studies the analysis should be done keeping participants in the group they were assigned to. The goals of randomisation are to avoid confounding and selection bias. If non-compliant participants or those receiving a wrong intervention are excluded from the analysis, the branches of a randomised trial may not be comparable. There are exceptions to the rule of intention to treat analysis.[67]

Authors' affiliations

M Delgado-Rodríguez, Division of Preventive Medicine and Public Health, University of Jaen, Spain
J Llorca, Department of Preventive Medicine and Public Health, University of Cantabria, Spain

REFERENCES

1 **Rothman K**, Greenland S. *Modern epidemiology*. 2nd ed. Boston: Lippicontt-Raven, 1998.
2 **Kleinbaum DG**, Kupper LL, Morgenstern H. *Epidemiologic research*. Belmont, CA: Lifetime Learning Publications, 1982.
3 **Sackett DL**. Bias in analytic research. *J Chron Dis* 1979;**32**:51–63.
4 **Choi BCK**. Bias, overview. In: Gail MH, Benichou J, eds. *Encyclopedia of epidemiologic methods*. Chichester: Wiley, 2000:74–82.
5 **Maclure M**, Schneeweiss S. Causation of bias: the episcope. *Epidemiology* 2001;**12**:114–22.
6 **Steineck G**, Ahlbom A. A definition of bias founded on the concept of the study base. *Epidemiology* 1992;**3**:477–82.
7 **Ellenberg JH**. Selection bias in observational and experimental studies. *Stat Med* 1994;**13**:557–67.
8 **Kleinbaum DG**, Morgenstern H, Kupper LL. Selection bias in epidemiologic studies. *Am J Epidemiol* 1981;**113**:452–63.
9 **Greenland S**. Basic methods for sensitivity analysis of biases. *Int J Epidemiol* 1996;**25**:1107–16.
10 **Pollock AM**, Benster R, Vickers N. Why did treatments rates for colorectal cancer in south east England fall between 1982 and 1988? The effect of case ascertainment and registration bias. *J Public Health Med* 1995;**17**:419–28.
11 **Chiang CL**. Competing risks in mortality analysis. *Annu Rev Public Health* 1991;**12**:281–307.
12 **Gail M**. Competing risks. In: Kotz S, Johnson NL, eds. *Encyclopedia of statistical sciences*. Vol 2. New York: Wiley, 1982:75–81.
13 **Simon R**. Length biased sampling in etiologic studies. *Am J Epidemiol* 1980;**111**:444–52.
14 **Neyman J**. Statistics: servant of all sciences. *Science* 1955;**122**:401–6.

15 **Hill G**, Connelly J, Hébert R, *et al.* Neyman's bias re-visited. *J Clin Epidemiol* 2003;**56**:293–6.
16 **Ransohoff DF**, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;**299**:926–30.
17 **Glesby MJ**, Hoover DR. Survivor treatment selection bias in observational studies: examples from the AIDS literature. *Ann Intern Med* 1996;**124**:999–1005.
18 **Arrighi HM**, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology* 1994;**5**:189–96.
19 **Berkson J**. Limitations of the application of fourfold table analysis to hospital data. *Biomet Bull* 1946;**2**:47–53.
20 **Feinstein AR**, Walter SD, Horwitz RI. An analysis of Berkson's bias in case-control studies. *J Chron Dis* 1986;**39**:495–504.
21 **Flanders WD**, Boyle CA, Boring JR. Bias associated with differential hospitalization rates in incident case-control studies. *J Clin Epidemiol* 1989;**42**:395–401.
22 **Horwitz RI**, Feinstein AR. Exclusion bias and the false relationship of reserpine and breast cancer. *Arch Intern Med* 1985;**145**:1873–5.
23 **Robins JM**, Pike M. The validity of case-control studies with nonrandom selection of controls. *Epidemiology* 1990;**1**:273–84.
24 **Goldstein AM**, Hodge SE, Haile RWC. Selection bias in case-control studies using relatives as controls. *Int J Epidemiol* 1989;**18**:985–9.
25 **Egger M**, Zellweger-Zähner, Schneider M, *et al.* Language bias in randomised controlled trials published in English and German. *Lancet* 1997;**350**:326–9.
26 **Jüni P**, Holenstein F, Sterne J, *et al.* Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol* 2002;**31**:115–23.
27 **Brick JM**, Waksberg J, Kulp D, *et al.* Bias in list-assisted telephone samples. *Public Opinion Quarterly* 1995;**59**:218–35.
28 **Callagham M**, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA* 2002;**287**:2847–50.
29 **Kjaergard LL**, Gluud C. Citation bias of hepato-biliary randomized clinical trials. *J Clin Epidemiol* 2002;**55**:407–10.
30 **Song F**, Easterwood A, Gilbody S, *et al.* Publication and related biases. *Health Technol Assess* 2000;**4**:1–115.
31 **Sterne JAC**, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000;**53**:1119–29.
32 **Frisch M**, Olsen JH, Bautz A, *et al.* Benign anal lesions and the risk of anal cancer. *N Engl J Med* 1994;**331**:300–2.
33 **Greenland S**. Response and follow-up bias in cohort studies. *Am J Epidemiol* 1977;**106**:184–7.
34 **Melton JL III** , Dyck PJ, Karnes JL, *et al.* Non-response bias in studies of diabetic complications: the Rochester diabetic neuropathy study. *J Clin Epidemiol* 1993;**46**:341–8.
35 **Froom P**, Melamed S, Kristal-Boneh E, *et al.* Healthy volunteer effect in industrial workers. *J Clin Epidemiol* 1999;**52**:731–5.
36 **Copeland KT**, Checkoway H, McMichael AJ, *et al.* Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977;**105**:488–95.
37 **Brenner H**, Blettner M. Misclassification bias arising from random error in exposure measurement: implications for dual measurement strategies. *Am J Epidemiol* 1993;**138**:453–61.
38 **Dosemeci M**, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect toward the null value. *Am J Epidemiol* 1990;**132**:746–8.
39 **Johannes CB**, Crawford SL, McKinlay JB. Interviewer effects in a cohort study. Results from the Massachusetts women's health study. *Am J Epidemiol* 1997;**146**:429–38.
40 **Neugebauer R**, Ng S. Differential recall as a source of bias in epidemiologic research. *J Clin Epidemiol* 1990;**43**:1337–41.
41 **Khoury MJ**, Flanders WD. Bias in using family history as a risk factor in case-control studies of disease. *Epidemiology* 1995;**6**:511–19.
42 **Wynder EL**. Investigator bias and interviewer bias: the problem of systematic error in epidemiology. *J Clin Epidemiol* 1994;**47**:825–7.
43 **Willet W**. *Nutritional epidemiology*. New York: Oxford University Press, 1989.
44 **Schulz KF**, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002;**359**:696–700.
45 **Morgenstern H**. Ecologic study. In: Gail MH, Benichou J, eds. *Encyclopedia of epidemiologic methods*. New York: Wiley, 2000:315–36.
46 **Davis CE**. The effect of regression to the mean in epidemiologic and clinical studies. *Am J Epidemiol* 1976;**104**:493–8.
47 **Bland JM**, Altman DG. Some examples of regression towards the mean. *BMJ* 1994;**309**:780.
48 **MacMahon S**, Peto R, Cutler J, *et al.* Blood pressure, stroke, and coronary heart disease. Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990;**335**:765–74.
49 **Wickström G**, Bendix T. The ''Hawthorne effect''—what did the original Hawthorne studies actually show? *Scand J Work Environ Health* 2000;**26**:363–7.
50 **Hernández Aguado I**, Ruiz Atienza L, Fernández García I, *et al.* Evaluación de la variabilidad del diagnóstico rápido de infección por *Chlamydia tracomatis* mediante inmunofluorescencia directa. *Med Clin (Barc)* 1990;**94**:9–11.
51 **Fletcher RH**, Fletcher SW, Wagner EH. *Clinical epidemiology. The essentials*. 2nd ed. Baltimore: Williams and Wilkins, 1988:161.
52 **Gullo L**, Pezzilli R, Morselli-Labate AM and the Italian Pancreatic Cancer Study Group. Diabetes and the risk of pancreatic cancer. *N Engl J Med* 1994;**331**:81–4.
53 **Feinstein AR**, Horwitz RI. An algebraic analysis of biases due to exclusion, susceptibility, and protophatic prescription in case-control research. *J Chron Dis* 1981;**34**:393–403.
54 **Shaper AG**, Walker M, Wannamethee G. Alcohol and mortality in British men: explaining the U-shaped curve. *Lancet* 1988;ii:1267–73.
55 **Feinstein AR**, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med* 1985;**312**:1604–8.
56 **Zhou XH**. Correcting for verification bias in studies of a diagnostics test's accuracy. *Stat Method Med Res* 1998;**7**:337–53.
57 **Knottnerus JA**, Leffers P. The influence of referral patterns on the characteristics of diagnostic tests. *J Clin Epidemiol* 1992;**45**:1143–54.
58 **Maldonado G**, Greenland S. Estimating causal effects. *Int J Epidemiol* 2002;**31**:422–9.
59 **Greenland S**, Brumback B. An overview of relations among causal modelling methods. *Int J Epidemiol* 2002;**31**:1030–7.
60 **Greenland S**. The effect of misclassification in the presence of covariates. *Am J Epidemiol* 1980;**112**:564–9.
61 **Brenner H**. Bias due to non-differential misclassification of polytomous confounders. *J Clin Epidemiol* 1993;**46**:57–63.
62 **Salas M**, Hofman A, Stricker BHC. Confounding by indication: an example of variation in the use of epidemiologic terminology. *Am J Epidemiol* 1999;**149**:981–3.
63 **Walker AM**. Confounding by indication. *Epidemiology* 1996;**7**:335–6.
64 **Moher D**, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;**357**:1191–4.
65 **Schulz KF**, Chalmers I, Hayes RJ, *et al.* Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.
66 **Murray DM**. Statistical models appropriate for designs often used in group-randomized trials. *Stat Med* 2001;**20**:1373–85.
67 **Lewis JA**, Machin D. Intention to treat—who should use ITT? *Br J Cancer* 1993;**68**:647–50.