

Facing the Facts

Investigative Radiology Must Face the Threat Presented by Systematic Errors



Prof. Werner Albert Golder

Correspondence

Prof. Werner Albert Golder
Radiologie, Association d'Imagerie Medicale, 65 rue Raymond
Poincaré, 10000 Troyes, France
Tel.: ++ 33/3/25 73 27 58
Fax: ++ 33/3/25 82 60 43
werner.golder@orange.fr

Bibliography

DOI <https://doi.org/10.1055/s-0043-123828>
Published online: 28.12.2017
Fortschr Röntgenstr 2018; 190: 235–236
© Georg Thieme Verlag KG, Stuttgart · New York
ISSN 1438-9029

After an initial test and introductory phase characterized by openness and optimism, many new diagnostic and therapeutic methods ultimately result in varying degrees of disappointment. The improvement and quality control of the clinical studies that are needed for the tests by the Institute for Quality and Efficiency in Health Care and the Federal Joint Committee and for the transition of innovations into reimbursable services in light of limited financial resources have not been able to slow, let alone stop, this trend. The fact that even studies designed and conducted according to strict criteria can yield positive results even though there is actually no positive effect or there are even disadvantages is often because the occurrence of systematic errors, referred to as a bias, is given insufficient consideration.

Systematic errors can negatively affect the conditions for method comparison, which are optimized in a clinical study, in that they produce results that deviate from the true values in a specific direction. As a result, differences are increased, decreased or even inverted in a targeted manner. Knowledge of the sources of the bias and the selection of suitable measures for decreasing their influence are of fundamental importance for clinical studies. Studies including patients and subjects are already associated with so many unknowns. Systematic errors should be avoided to the greatest extent possible. The case-control, cohort, and interventional studies predominantly seen in radiology are susceptible to numerous systematic errors and it is difficult to safeguard against them. Knowledge of the most important types of bias is absolutely essential.

Despite careful planning and implementation, systematic errors cannot be categorically prevented in clinical studies [1]. However, efforts to reduce the number of errors are a central quality feature and help to determine the reliability of results. If a bias cannot be ruled out by the method, its potential effects on results must be determined as precisely as possible. This requires

targeted identification and determination of the direction of the error. However, while other variables such as chance and interference can be quantified and subsequently mitigated by mathematical corrections, there is no such option for systematic errors. Systematic errors also cannot be eliminated by the increasing the sample size. This concept is suitable only for reducing random errors. Systematic errors must be addressed in the planning phase of a study. Randomization and blinding are the most important instruments for preventing these errors. Randomization promotes bias-free distribution of known as well as unknown influencing variables between the different groups. Blinding reduces systematic errors resulting from knowledge of the participants' history.

- One reason for the high values for statistical accuracy documented in many publications is the fact that the analysis is performed by specialists/subspecialists and not by physicians primarily performing the general duties of their specialty. This bias can be reduced by the measurement data being evaluated in parallel by experts who are not directly connected to the research project.
- The homogeneous composition of study collectives also has disadvantages. Studies comparing patients with full-blown disease and healthy controls are less representative of the spectrum of patients seen in practice than studies also including patients with fewer symptoms and controls with other diseases, some of which can be mistaken for the disease being studied.
- High-quality studies apply high standards with respect to the selection and provision of the reference standard for all participants. Ideally, the examiner performing the test procedure is blinded to the results of the reference standard test and vice versa. However, for practical reasons, a less suitable parameter must at times be used for the reference standard and some-

times even this parameter is not uniformly recorded for all participants and applied to the analysis. Such a source of error must be taken into consideration in all comparative studies of diagnostic tests.

- Many examiners believe that the ideal study design entails disease verification according to strict criteria. However, studies in which cases that cannot be definitively confirmed are ruled out are subject to a systematic test error. Only subjecting patients with definitely positive test results to the reference standard test can result in an overestimation of the sensitivity.
- In a group of patients with the same diagnosis, identical findings are formulated more liberally than in a heterogeneous population. To avoid distortion of the determined accuracy of a test, examiners should always use samples in which the prevalence of the symptom or the disease corresponds to that of the clinically relevant population. However, in practice it is often difficult to meet this requirement. However, the large number of similar images and issues seen when evaluating radiological studies is not representative of the prevalence of such materials in daily health care even in specialized facilities.
- The influence of learning curves on the practical application of new techniques distorts results both on the part of study physicians and test persons throughout the entire observation period. Two mechanisms are important here. On the one hand, in the comparison of two methods by a single person, the new currently researched method has an advantage over the control method if beginner's mistakes and minor errors are tolerated. On the other hand, performance can be underestimated when hardware and software capabilities are not fully utilized due to a lack of knowledge or hesitation.
- The more sensitive a test is, the longer the population on which it was performed seems to live. Accelerating the time of diagnosis distorts above all the results of studies regarding the effectiveness and efficiency of early detection tests for malignant tumors. Survival times of tested and untested persons differ by more than the lead time only if the test and the resulting measures actually extend survival time. Otherwise, the earlier diagnosis only results in tested persons knowing earlier that they are sick but not in an extension of survival time beyond the lead time.

- New diagnostic/therapeutic methods are typically tested by representatives of disciplines that then subsequently use them with primary responsibility. However, this privilege bears the fundamental risk that the evaluation of the method is too positive out of self-interest. However, negative or cautionary voices can prevail when representatives of one or more competitive disciplines perform the test. Distortions due to such professional biases can be reduced when studies are performed on an interdisciplinary basis or at least the analysis is performed in an interdisciplinary team.

Even with careful planning, many systematic errors that jeopardize clinical studies can only be reduced, not eliminated. As a result, it is extremely important to analyze distortions in detail and to fully document acquired results so that the insights that can be gained from studies can be used as the basis for reimbursable medical services. Biases must be separately and comprehensively examined and shown for every scientific study. Multiple biases usually add up, resulting in a completely distorted picture. Analysis of the role of bias as an alternative explanation for an observed connection is essential in the interpretation of every study result even if this decreases the scientific yield. Studies that are largely free of systematic distortions almost always yield results that are less statistically significant than studies that are subject to bias. Examiners should describe all potential distortions of their work and the measures taken to eliminate these in order to provide a comprehensive overview of the possible influence of systematic errors on findings and conclusions [2]. This analysis should be an independent section of the discussion section of every scientific publication and should be requested by editors upon submission as an obligatory part of a manuscript. Allusions and brief summaries are not sufficient. Studies not including a sufficient explanation regarding bias should not be accepted for publication.

References

- [1] Sackett DL. Bias in analytic research. *J Chronic Dis* 1979; 32: 51–63
- [2] Bossuyt PM, Reitsma JB, Bruns DE et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003; 138: W1–W12