III. Causality Studies

Juan O. Talavera, Niels H. Wacher-Rodarte, Rodolfo Rivas-Ruiz

Although the need of solving a clinical problem leads to the establishment of a starting point for approaching it (risk, prognosis or treatment study), in all cases, there is an attempt to attribute causality. Clinical reasoning, analyzed in detail in the book *Clinical Epidemiology. The architecture of clinical research* offers a simple guideline for undestanding this phenomenon and uses three components: baseline state, maneuver and outcome. In this model, different systematic errors are described (biases), which can occur when features of these basic components are overlooked. Omisions of characteristics at the baseline state produce an *inadequate assembly* of the population and the *susceptibility bias*; in the application or assessment of the maneuver, the *execution bias*; and in the assessment of the outcome, the *detection bias* and the *transference bias*. Thus, it is important to emphasize that if this form of reasoning facilitates the comprehension of the causal phenomenon, variables to be selected in studies where causality will be attributed or not to them require additional clinical reasonings assessing their relevance.

Key words
research
causality

This article was originally published in Rev Med Inst Med Seguro Soc 2011; 49 (3): 289-294 and it has been reviewed for this issue.

Introduction

When trying to predict a future event, the physician has to differentiate two processes: one that occurs before the onset of the disease and other that develops once the disease is present. The first is known as risk and it is characterized by the association between a series of factors present in the healthy subject (known as risk factors) and the development of the disease; the second is known as prognosis and it is characterized by the association between a series of features present at the beginning of the disease (known as prognostic indicators) and its outcome.

Multiple interventions, either preventive or therapeutic, add up to these two events; the former are intended to prevent the onset of the disease and the latter, to revert or reduce the damage caused by it. The event whereby a baseline condition (health or disease) is modified by a maneuver (risk factors, prognostic indicators or treatment), and which in turn produces a new condition known as outcome (prevention or onset of the disease and progression or resolution of harm), corresponds to a causative event. That is, in these three cases –whether our objective consists in identifying risk factors, an etiologic agent, prognostic indicators or assessing a treatment– attribution of causality is intended.

Although the need to solve a clinical problem leads us to establish a starting point to address it –risk, prognosis or treatment study–, in the real world there is a strong association between its components. For this reason, when assessing any of them, it is essential for the relevance of the other two to be considered within the assessment. This action is often carried out under the term *control of confounding factors*.

Thus, the study of causality for assessing a treatment is not only limited to the evaluation of therapy, but it obliges to estimate the contribution of all prognostic indicators existing at baseline state that participate in the disease of interest.

Likewise, when trying to prevent the onset of a disease with some maneuver, we must assess the different risk factors specifically associated with this disease. This requirement of measuring the impact of the different risk factors and prognostic indicators when assessing a therapy is consistent with the requirement of assessing the different therapeutic procedures when what we are trying to evaluate are the risk factors or prognostic indicators.

Clinical Reasoning in Causality Studies

Clinical reasoning, which is analyzed in detail in the book *Clinical Epidemiology. The architecture of clinical research* offers a simple approach for understanding

the phenomenon of causality. Figure 1 shows the basic model comprising the baseline state, the maneuver and the outcome. This model describes different systematic errors (biases) that may contribute to the omission of some characteristics of the three basic components.

Errors at the Baseline State

The first two errors are related with omissions of baseline state characteristics and these are improper assembly and susceptibility bias.

Improper assembly refers to the selection of a population not susceptible to experience the outcome of interest with a proposed maneuver; for example, it is rather impractical to test a vaccine in a population with low incidence of the disease we are trying to prevent, since the size of the sample would have to be enormous; it is also inconvenient to assess the kidney-protecting effect of an ACE in a population of newly-diagnosed diabetic patients, since the follow-up would have to be very long.

Susceptibility bias refers to the pre-maneuver likelihood that the subject has of experiencing a certain outcome; for example, the presence of overweight or obesity increases the likelihood of an infarction in a diabetic patient, regardless of the poor metabolic control he may have.

The characteristics that must describe the baseline state to avoid these errors are shown in Figures 2a and 2b, i.e., the method used to select the population, the diagnostic demarcation and the prognostic stratification.

Within the prognostic stratification, anatomo-histology has been used as the main indicator, especially in oncology, followed by the functional aspect. In clinical practice, it is common to use multiple prognostic indicators in order to stage the disease according to the patient's condition. The following stratification groupings are the most common:

Primary

- Stratification by status: it includes the performance, nutritional and mental status of the patient. Performance status has ben assessed with scales such as Karnovsky or ECOG, based on the patient's ability to perform his/her daily activities, in such a way that a patient who is not self-sufficient is more affected than that who can perform his/her tasks. Nutritional status impacts on the immune response and the hemodynamic stability. Patients with low albumin levels have been observed to show an important increase in mortality compared with those with higher levels. Other forms to assess nutritional status could be the body mass index and the waist-hip ratio when trying to assess the impact of overweight or body fat distribution; additionally, two of the most important features for assessing the mental status are the presence of depression and anxiety, among many other conditions.
- Morphologic stratification: it refers to the distinct location and damage of the pathology. An example is the histologic lineage of tumors and cytogenetic or immunophenotypical markers (for example, two tumors with the same extent of disease may have different prognosis according the histologic lineage, the presence of tumor markers or karyotype alterations; also, a patient with heart failure may have different prognosis according to the degree and type of valvular damage).
- Clinical stratification: it considers the severity of the disease, for example, the patient with grade IV heart failure (acute pulmonary edema) does not have the same probability of death than the patient with grade II (dyspnea with moderate exertion), even when the anatomical condition in both cases may be a mitral stenosis with the same valvular opening diameter.
- Chronometric stratification: it considers two components, the patients' age and the length of the disease. Regarding the first one, many diseases have

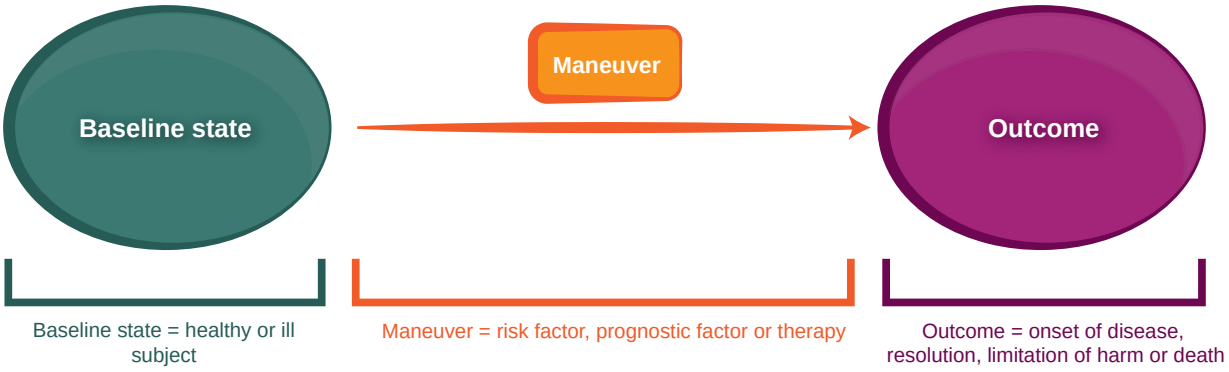

Figure 1 Basic model of the causality phenomenon

greater impact at both extremes of life and are associated with higher susceptibility to a poor outcome; additionally, older individuals have lower life expectancy. Regarding the length of the disease, if two patients suffer the same harm, but in one of them the disease is of recent onset while in the other it is of long evolution, the prognosis will be better in the latter since those patients with less aggressive disease have already been selected.

- Stratification by comorbidity: it refers to the coexistence of any other pathological process that may alter the result of interest. Different conditions exert different impact on the outcome, and even in a same condition, the impact is generally related with the degree of illness; for example, in a patient with acute myocardial infarction, the prognosis is better when the comorbidity is rheumatoid arthritis than when it is diabetes mellitus.
- Stratification by previous maneuver: two items can be identified here: the first and most widely used is the early response to a preventive or therapeutic maneuver, i.e., a better prognosis is expected upon an early favorable response. The second refers to the adverse impact of a maneuver. Practically every maneuver is known to entail a risk; however, not in all of them it has the same magnitude. Thus, safety should be considered as a prognostic indicator for any therapy.

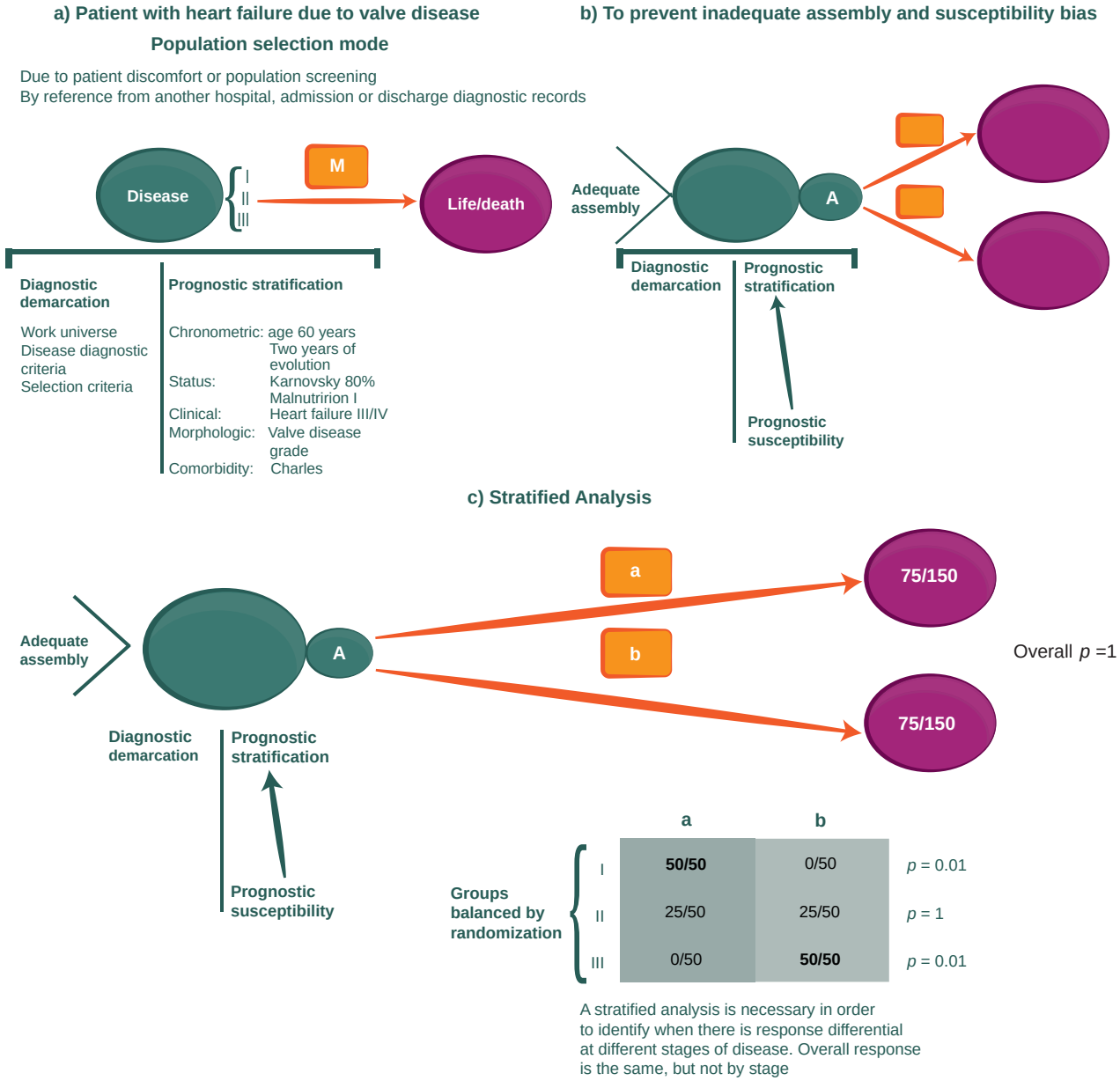

Figure 2 Features to be considered at the baseline state

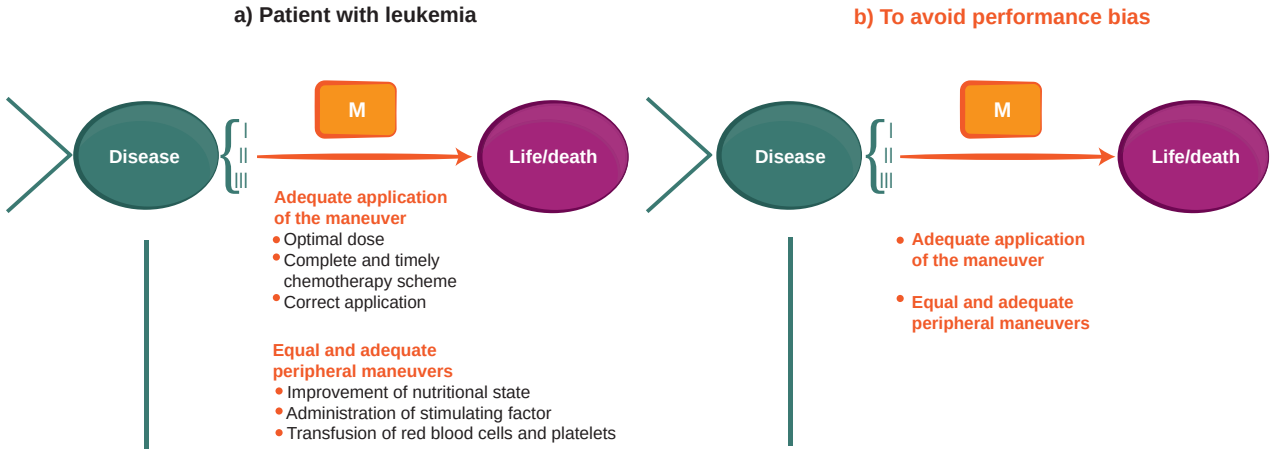

Figure 3 Features to consider in the maneuver

- Stratification by inheritance: the impact of genetic makeup has been identified as a risk factor for several diseases and with an increased aggressiveness thereof or higher risk of harm to target organs, as in diabetes.

Secondary

- Social, economic and cultural conditions, as well as the ways of coping with disease, often have a lower impact than the biological components within the prognosis; however, sometimes they are crucial, such as having access to health care services in emergency events, or the change in lifestyle in some chronic diseases.

A distinctive strategy of clinical trials to avoid susceptibility bias is the random allocation of subjects to the treatment arm, seeking, among other things, that known and unknown factors potentially related with the outcome are evenly distributed between the groups to be compared. Other benefit is to prevent that those in charge the allocation are tempted to include a subject with better prognosis in a particular arm, since randomization facilitates the blinding of treatments and seeks to homogeneously distribute the subjects with different likelihood of treatment adherence and different likelihood of study dropout. It should remain clear that although random allocation seeks that the groups to be compared are homogeneously distributed at their baseline state, it does not show the effect of the maneuvers on the different strata (Figure 2c).

Errors in the Maneuver

The third systematic error, known as performance bias, is related with omissions in the application or assessment of the maneuver, and it refers to the differences generated by quality differences between the maneuvers to be compared or by an uneven use of additional maneuvers between groups (also known as peripheral maneuvers); for example, a surgery is not the same when performed by a recently graduated surgeon than when performed by a physician with extensive experience, nor are comparable two surgeries when in one of them the patients are well nourished or brought to hemoglobin normal values, while in the other group they are not. Features that have to be considered in the maneuvers in order to prevent these errors are shown in Figures 3a and 3b, which consist in adequate application of the maneuver and equal application of peripheral maneuvers.

In clinical trials, there is a strategy intended to handle errors generated by an inadequate application of the maneuver, which is the way of analyzing the information, either by means of an intention-to-treat analysis or a per-protocol analysis. The *intent-to treat analysis* consists in analyzing the subjects in the group they were allocated to at the beginning of the study, regardless if they were compliant with the therapeutic protocol or not. The *per-protocol analysis* consists in analyzing only those subjects who were compliant with the therapeutic protocol. In observational studies, since there is no randomization to the maneuver, this is graded within the groups, thus enabling the comparison of the different degrees of quality in the maneuver application.

Errors in the Outcome

Detection bias occurs during the assessment of the outcome, which relates to an uneven detection of the outcome between groups and it occurs mainly for two reasons:

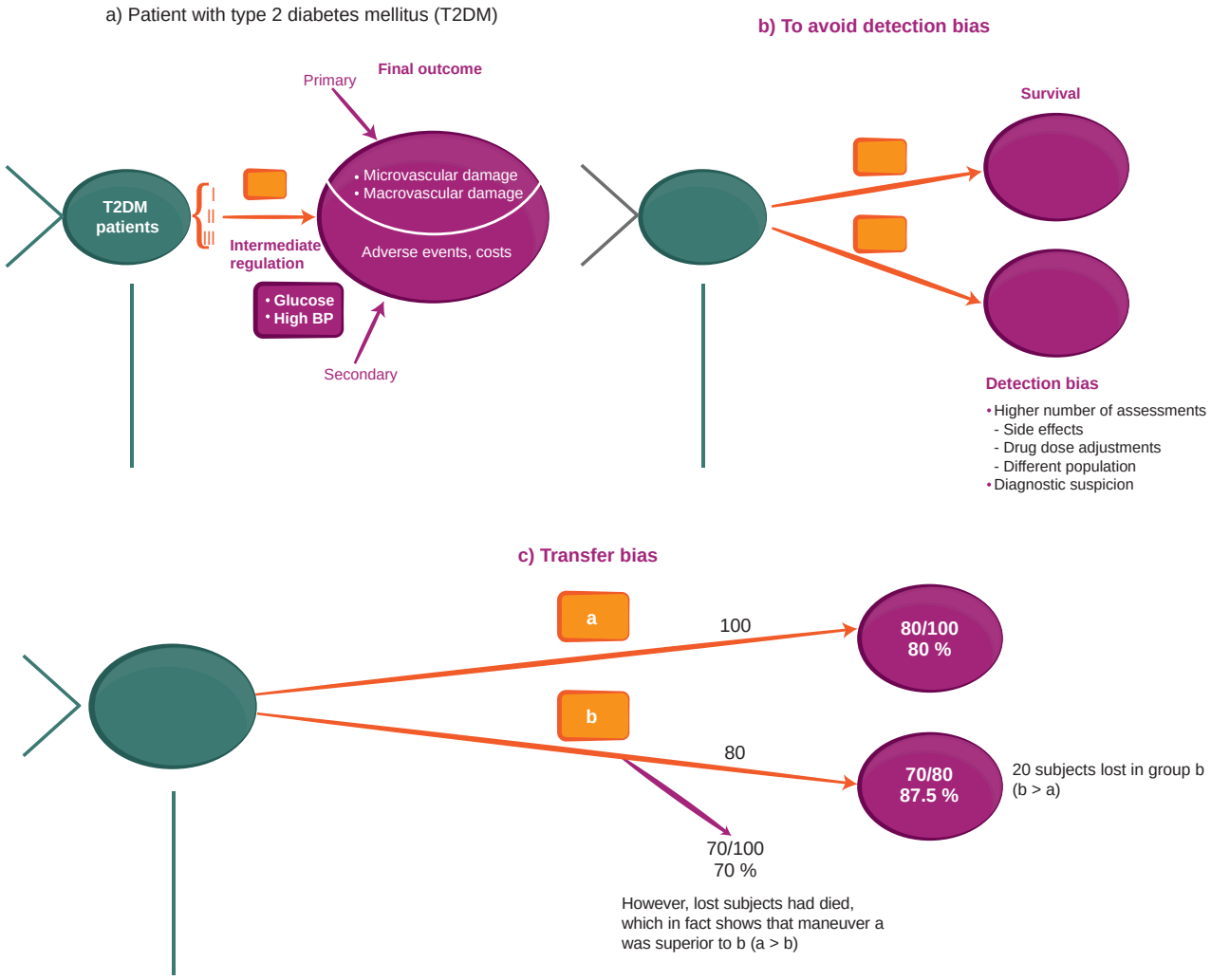

Figure 4 Main features to consider when assessing the outcome

- A higher number of assessments in some group, mainly due to more side effects, continuous dose adjustments or comparison of populations with different healthcare accessibility.
- Presence of diagnostic suspicion.

In the assessment of the outcome it is important to identify whether it is a final outcome or an intermediate regulation; for example, in the diabetic patient, the final outcome is to prevent damage in target organs; however, an intermediate regulation is glucose control; the latter may be considered a final outcome if symptomatology is trying to be reduced in the uncontrolled patient.

Another important aspect in outcome assessment is the identification and differentiation between the primary and the secondary outcome. This point is relevant since the selection criteria and the prognostic stratification, as well as the maneuver and the sample size estimation are carried out on the primary outcome and not on the secondary. Therefore, the results obtained in most studies are only exploratory for secondary outcomes (Figures 4a and 4b).

The last bias is also related with the outcome; it is generated by the loss of subjects under study and it is known as transfer bias (Figure 4c). Although in prospective studies the sample size is increased by 20 % in order to account for potential withdrawals, it is important to emphasize that this increase does not solve the transfer bias, but it rather maintains the stability of the data.

Final Considerations

In longitudinal studies, it is easy to apply these guidelines to study the phenomenon of causality; in the trans-

versal ones they continue to be applicable, but this is a major challenge that translates into the creation of an artificial model regarding the temporary establishment of its components. Taking into account the elements described herein is recommended, not only for the reading of a causality study, but also for the creation of a research proposal.

It is important to emphasize that if this form of reasoning facilitates the understanding of the causative phenomenon, the appropriate thing to do for selecting those variables to which causality will be attributed to or not, is taking into account additional clinical considerations assessing their relevance. The basic principles were described in 1965 by Sir Austin Bradford Hill and were updated in 2000 by Kaufman and Poole; surely, over time, the number of factors to consider when judging a potential causal relationship will increase.

We hope that the causality approach herein described, which breaks down the basis of clinical practice, will facilitate the interpretation of medical literature and serve as guidance for the planning of research proposals and to increase the quality of medical care.

Bibliography

1. Charlson ME, Frederic LS. The therapeutic efficacy of critical care units from two perspectives; a traditional cohort approach vs. a new case control methodology. J Chron Dis. 1987;40:31-39.
2. Feinstein AR. Clinical epidemiology. The architecture of clinical research. Philadelphia: WB Saunders; 1985.
3. Feinstein AR. Directionality and scientific Inference. J Clin Epidemiol. 1989;42:829-833.
4. Fletcher R, Fletcher S, Wagner E. Clinical epidemiolgy: the essentials. 2nd. ed. Baltimore: Williams & Wilkins; 1988.
5. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. Baltimore: Williams & Wilkins; 2008.
6. Sackett D, Haynes R, Tugwell P. Epidemiología clínica una ciencia básica para la medicina clínica. Madrid: Ediciones Díaz de Santos; 1989.
7. Talavera JO. Pronóstico. En: Ramiro M, Lifshitz A, Halabe J, Frati A, editores. El internista. Medicina interna para internistas. 3a. ed. México: Nieto Editores; 2008. p. 1893-1898.