Assessing the diagnostic ability of medical tests with status defined by right-censored data at a specific time t

Sara Perez-Jaume[1,2], *Jaume Mora*[2], *Josep L Carrasco*[1]

[1]Biostatistics Unit, Department of Basic Clinical Practice, University of Barcelona (SPJ: saraperez@ub.edu, JLC: jlcarrasco@ub.edu)

[2]Sant Joan de Déu Pediatric Cancer Center Barcelona, Institut de Recerca Sant Joan de Déu, Barcelona (SPJ: sara.perez@sjd.es, JM: jaume.mora@sjd.es)

A diagnostic test is a tool applied to individuals with the objective of diagnosing a certain condition or status of health or prognosticating the occurrence of a particular event. Diagnostic tests are usually used to distinguish between diseased and non-diseased subjects and they are of the utmost importance in clinical and biomedical research, since the decisions made on the basis of a diagnostic test might have crucial implications in the management of the patients. The most usual setting considers two possible statuses for the subjects, that is, the outcome of interest is binary. When assessing the diagnostic ability of binary tests, common measures of accuracy are sensitivity and specificity. In the case of continuous diagnostic tests, the area under the receiver operating characteristic (ROC) curve (AUC) is a commonly used index to evaluate the ability of the test to discriminate among the two true states. Moreover, it is usual to choose a threshold that is optimal in some sense to define the test results as positive or negative. One of the most popular criteria to estimate optimal thresholds is the maximisation of the Youden index, which is defined as the sum of sensitivity and specificity diminished in one unit. Once a threshold has been established, a binary test can be constructed from the continuous one and measures such as sensitivity and specificity might also be estimated to assess the diagnostic ability of the test.

Sometimes, the interest lies in predicting the binary status of the subjects at a certain pre-specified time-point t, which is of interest for clinical or biological reasons. When this status is available for all individuals, sensitivity, specificity, AUC and optimal thresholds can be estimated using the standard methodology for the two-state setting. The problem arises when the status is defined by an event occurring at t or before. If the time to event is right-censored (that is, the event time is larger than the censoring time), the status at t is not observed for those patients censored before t. We consider this situation as a missing data problem, and thus we propose two approaches to deal with it. First, we introduce a simple method based on excluding those individuals with missing status at t. Then, we propose a more complex method that uses multiple imputation to recover knowledge about the status of those subjects at time t, using the information provided by the diagnostic test. Those methods are used for estimation and inference about sensitivity, specificity, AUC and optimal thresholds. We illustrate these methodologies by applying them to real case examples about cancer patients. The simulation study demonstrated low biases with both methods and showed the advantages of the imputation method when compared to the missing exclusion approach, especially with small or moderate sample sizes and in scenarios with high proportions of missingness.

Keywords: diagnostic ability assessment, missing data imputation, right-censored data