



Title	Evaluation of Diagnostic Tests When There Is no Gold Standard. A Review of Methods
Agency	NETSCC, HTA, NIHR Evaluation and Trials Coordinating Centre Alpha House, University of Southampton Science Park, Southampton, SO16 7NS, United Kingdom; Tel: +44 2380 595 586, Fax: +44 2380 595 639; hta@soton.ac.uk, www.hta.ac.uk
Reference	Volume 11.50, ISSN 1366-5278. www.hta.ac.uk/project/1573.asp

Aim

To generate a classification of methods to evaluate medical tests in the absence of a gold standard.

Conclusions and results

Most methods try to impute, adjust, or construct a reference standard in an effort to obtain familiar diagnostic accuracy statistics, eg, sensitivity and specificity. In situations that deviate only marginally from the classical diagnostic accuracy paradigm, these are valuable methods. However, in situations where an acceptable reference standard does not exist, applying the concept of clinical test validation can provide a significant methodological advance. All methods summarized in this report need further development. Some methods, such as the construction of a reference standard using panel consensus methods and validation of tests outside the accuracy paradigm, are particularly promising, but are lacking in methodological research. These methods deserve particular attention in future research. Available methods were classified into 4 main groups. The first method group (impute or adjust for missing data on reference standard) pays careful attention to the pattern and fraction of missing values. The second group (correct imperfect reference standard) can be useful if there is reliable information about the degree of imperfection of the reference standard and about the correlation of the errors between the index test and the reference standard. The third method group (construct reference standard) combines multiple test results to construct a reference standard outcome including deterministic predefined rules, consensus procedures, and statistical modeling (latent class analysis). In the final group (validate index test results) the diagnostic test accuracy paradigm is abandoned and the research uses different methods to examine whether the results of an index test are meaningful in practice, eg, by relating index test results to other relevant clinical characteristics and future clinical events.

Recommendations

In situations where an acceptable reference standard does not exist, holding on to the accuracy paradigm is less fruitful. In these situations, applying the concept of clinical test validation can provide a significant methodological advance. Validating a test means that scientists and practitioners examine, using several different methods, whether the results of an index test are meaningful in practice. Validation will always be a gradual process. It will involve the scientific and clinical community defining a threshold, a point in the validation process, whereby the information gathered would be considered sufficient to allow clinical use of the test with confidence.

Methods

See Executive Summary link at www.hta.ac.uk/project/1573.asp.

Further research/reviews required

All methods summarized in this report need further development. Some methods, eg, construction of a reference standard using panel consensus methods and validation of tests outside the accuracy paradigm, are particularly promising, but are lacking in methodological research. These methods deserve particular attention in future research.