Diagnostic research in perspective: examples of retrieval, synthesis and analysis
Bachmann, L.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Introduction
Chapter 1

Clinicians are expected to make a timely diagnosis, and to favourably influence their patients' course with appropriate treatment or, alternatively, withholding unnecessary treatment. A wrong diagnosis can harm patients either by exposing them to inappropriate therapy or denying them useful treatment, while correct diagnosis may allow timely use of effective treatment.

The methodological concept behind the bulk of diagnostic research reaches back to the late nineteen-fifties when a mathematician and a radiologist published an analysis about "the complicated reasoning processes inherited in medical diagnosis". This paper was perhaps most influential in establishing the prestige of assessing the accuracy of a diagnostic test. In principle this paper postulated that the parameters sensitivity and specificity must be viewed as characteristic properties of the diagnostic procedure. Once they are known, one only needs an estimate or an educated guess of the prevalence of the disease to calculate the predictive values of Bayes theorem. If more tests are available, the predictive value of a sequence of test results can be computed under the assumption of conditional independence, sometimes referred to as "Idiot's Bayes". This paradigm, yet statistically unsound, was so attractive that it became one of the basic concepts of medical decision-making and clinical epidemiology.

To assist clinicians in deriving posttest probabilities Fagan published a nomogram that allows the estimation of a disease probability using the pretest probability of disease and the test accuracy. This nomogram, has become an essential tool for evidence-based diagnosis, printed on coloured cards and disseminated through EBM textbooks.

Starting in the late seventies, alternative concepts of diagnostic research with roots in statistics have been proposed. Only recently these ideas have been discussed again more seriously in the medico-scientific community. In principle, this alternative viewpoint starts from the observation, that in clinical practice diagnosis is a process that integrates the information of a set of patient information, signs, symptoms clinical examination and other tests. Secondly, it takes into account that in clinical practice the health care provider is not interested in the likelihood of a positive test result (as it is provided in test accuracy research) but in the probability of disease of his patient. It was proposed that diagnostic research should also evaluate which set and sequence of diagnostic indicators is most efficient to estimate the probability of disease.

Evidence-based diagnosis

Central to the idea of evidence-based medicine is the identification of medical information for clinical practice. The systematic identification and summarisation of the available evidence have become a central element of research and informed decision-making in medicine. Historically, evidence-based medicine has focused on the summarisation of the evidence of therapeutic studies. Only recently, it has been recognised, that informed decision-making also clearly includes evidence-based diagnosis. The identification of diagnostic studies in biomedical databases is seen as the starting point of the summarisation process of a systematic review. Compared to literature retrieval of therapeutic trials, identification of diagnostic studies
is still recognised to be more difficult. The key argument raised here is the inconsistent indexing of diagnostic studies in biomedical databases. Several search filters have been published to assist clinicians or reviewers in their search activities/searches. All filters however still struggle with a low retrieval precision if retrieval sensitivity is to be high. Comprehensive searches therefore usually identify many irrelevant studies that researchers then have to screen in order to exclude them.

As outlined above, the vast majority of diagnostic studies assess the accuracy or characteristics of a test. Generally, study-level diagnostic systematic reviews therefore aim to summarise the available evidence on a single test in the diagnosis of a particular target disease. The research question usually contains as components the characteristics of the population under study, the test of interest and the definite diagnosis of the condition or disease. Results of such systematic reviews are presented as pooled sensitivities, specificities, predictive values, likelihood ratios or summary ROC curves.

In the light of the viable concerns raised both by statisticians and cognitive psychologists that test accuracy measures are prone to biases and are of limited use in clinical practice, it has been proposed that results of diagnostic accuracy reviews should be linked to results of therapeutic interventions. Consequently, the usefulness of a diagnostic test would be expressed as the comparison of the value of an efficacy measure, such as the number needed to treat, given a positive test result and the value of that measure if no testing is performed. These measures however are rarely seen in the literature.

Probability analysis

Thinking of diagnosis as the process of disease probability estimations using a sequence of tests, clinicians should have access to diagnostic measures of all tests applied in the diagnostic sequence. A well-formulated diagnostic research question could read something like this. Given a diagnostic suspicion that disease X may underlie this patient's complaints, which tests should I order to raise or lower my suspicion such that my treatment decision(s) would be different from the scenario in which I can order no test, and in which order should I apply the tests, if more than one were indicated.

Diagnostic research studies that could provide the answers to questions of the above type would perform, in each patient, all tests that are currently available for the diagnosis of X or a relevant subset thereof including demographics and other information of the clinical history. The (multivariable) analysis of this study would take into account redundancy of information between all the tests of the sequence. In addition, it would sort out the optimal sequence in which (some of) the more advanced tests should be ordered for patients with different profiles on the easily obtainable information from demographics, patient history, and, perhaps, physical examination. The results of such an analysis would be communicated to physicians as simple post-test-sequence probabilities of disease presence for specific patient profiles. For example, women above a certain age and with a specific ethnic background may need a different set of tests to establish a final diagnosis than other women. Or, they may be more (cost-) efficiently categorised by using a different order of testing.
Chapter 1

There is currently a dearth of such studies. This is possibly due to several factors. Firstly, these studies tend to be more expensive than test accuracy studies because all tests of the sequence have to be applied in all patients entering the study. In addition, dependent on the prevalence of the disease, these studies need to have a larger number of patients to derive valid and precise probability estimations. Usually these studies tend to conclude that further research on larger patient samples is required to make reliable inferences for practice.

Outline of the Thesis

This thesis contains examples of original research around all the issues of diagnostic research, which were addressed above. Chapter two and three consist of two studies that describe the construction of two methodological search filters to identify diagnostic studies in the two most frequently searched biomedical databases (Medline, Embase). Chapter four and five give two examples of systematic reviews of test accuracy studies. One review deals with a test used in diagnosis or screening (cervicovaginal fetal fibronectin to predict preterm birth) and the other summarises the evidence on the Ottawa ankle rule, a clinical decision aid or triage test invented to reduce the number of unnecessary radiographics in patients with ankle sprains. Both reviews aim to apply recent methodological propositions in the conduct of systematic reviews. Chapter six highlights the difficulties of understanding of test accuracy measures in clinical practice and discusses possible improvements. Chapter seven and eight, finally, present two examples of multivariable probability analysis. Chapter nine discusses the presented studies and reflects on possible future developments of diagnostic research.

References


