Diagnostic research

Challenges in dia-prognostic research

J A Knottnerus

The relevance of design issues and new methods

It is sad to comment on work of a founding father of clinical epidemiology, who so profoundly enjoyed the scientific debate, without him being able to respond. But it is not only sad, it is also a great honour to have the opportunity to discuss one of the last pieces of this great architect of clinical research.[1]

The points made by Alvan Feinstein are clear and I agree with most of his analysis, but some additional challenges should be emphasised.

LIMITATIONS OF "NOSOLOGICAL INDICES"

Indeed, apart from interpretative problems, "nosological indices" such as sensitivity and specificity are problematic as they will vary with the spectrum and selection of patients.[2–4] However, also "diagnostic indices" (predictive accuracy) are not constants. Taking account of all relevant covariables is, if at all useful, generally impossible because many of these are not clear, unmeasurable, or too particularistic for a useful general evaluation. So, we must not overemphasise the value of diagnostic indices. And where we use them, we must be very specific as to objective, context, and setting. We should also be keen on setting specific external validation when we want to generalise observed index values for clinical purposes. Acknowledging all this, it cannot be denied that for clinical and epidemiological purposes there is often a need to use "nosological indices" to communicate general characteristics of a test. Also Feinstein did not escape from speaking about characteristics required for a "rule out" test (high sensitivity) compared with a "rule in" test (high specificity). The question what index is "the best" is too general, and therefore a non-issue. While in principle all indices (and combinatory roles, using multivariable analysis) can be calculated from data on the association between test outcome and disease, which index is actually used depends on objective and context. For example, knowing about sensitivity and specificity is important for the selection of the most appropriate screening test, while a high predictive value is important when evaluating the diagnostic value of a pre-mastectomy cancer diagnosis.

Bayes's theorem, ROC curves with varying cut off values, likelihood ratios, and logistic regression can be unattractive for many clinicians and also these entities are not insensitive for spectrum and selection.[5] But for various purposes they cannot be missed, for example, in data analysis, computer algorithms, and in finding optimal cut off values for screening tests. At the same time, many challenges in the context of data analysis remain to be adequately dealt with (such as: assessing additional diagnostic value of a test rather than accuracy in itself[6]; differences between multivariable analysis for diagnostic and aetiological purposes (with major emphasis on subgroup analysis for the former); meta-analysis of research data and making them tailormade at the same time; and bridging the gap between clinical reasoning in practice and products of data analysis.

DESIGN ISSUES MORE IMPORTANT

We should not forget that mathematical indices are just ways to summarise collected research data. For the quality of research, defining the research question, and methodological challenges in study design, are far more important. The high degree of difficulty of diagnostic research studies is reflected by various overviews over the past decades, repeatedly showing important flaws in a high proportion of studies,[7–9] and indicating only slow improvement.

Diagnostic research should be improved and refined with respect to strategy (where phase I to phase IV studies should be subsequently designed[1 10]), spectrum and selection effects, diagnostic and prognostic reference standards, and the assessment of the clinical impact of testing. Better methods to warrant and evaluate external clinical validity are required. Furthermore, proposed innovations must always be compared with achievements of experienced clinicians, before they are recommended as better. We need more understanding of the "doctor's black box" of diagnostic decision making, jumping between observations on groups and caring for individuals, including reassurance. This will facilitate more efficient diagnostic reasoning, and developing custom-made support systems. Efficiency in evaluation of the impact of diagnostic procedures can be gained if new data on specific aspects (for example, a diagnostic test) can be inserted in the mozaic of available evidence on a clinical problem, rather than studying the whole problem again whenever one element has changed. For this purpose, flexible scenario models of current clinical knowledge are needed. This can be strongly supported by ongoing prospective systematic review and meta-analysis of diagnostic studies and developing clinical databases.

DIA-PROGNOSTIC SIGNIFICANCE

As Feinstein points out, significance for prognosis and medical decision making is the justification of diagnosis. This insight is already often implemented and evaluated, for example in the process of staging tumours or in the clinical classification of heart failure. But we must recognise that (evaluation of) prognostic significance needs much more attention, especially where test technology advances without concomitant therapeutic improvements. A well known example is detailed, three dimensional MRI observations of the brain of which clinical relevance is not (yet) clear.

Innovation of biomedical knowledge and understanding pathophysiological processes is decisive for the development of tests with better "dia-prognostic" impact. DNA diagnostics is in fact prognostics as it touches the basic functional level of our biomedical nature. DNA testing will not only be supportive of genetic counselling and reproductive applications, but also for diagnostic and prognostic purposes. Moreover, DNA testing is expected to improve the targeting and dosing of interventions. Much work in this field is being done, for example, in cardiovascular medicine, oncology, and psychiatry. However, it will take much time before these promises will have impact in daily patient care. Considerable efforts are still needed not only in the laboratory but also in clinical epidemiological reseach[11] including long term follow up to clinically validate diagnostic and prognostic predictions.

INDIVIDUALISATION AND DEALING WITH (UN)CERTAINTY

In view of the ambition to develop a more tailormade, perhaps even individualised, "dia-prognostic" process, study population oriented validations will get increasingly under pressure. In this context, n=1 research, focused on individual patients, represents great methodological challenges. In addition, the moment of testing and the occurrence of a clinical problem will not necessarily be always related. The DNA profile with its predictive potentials is there from the beginning. Ethical questions on when to test, how to deal with privacy of genetic information, and the right of (not) knowing have to be addressed. Doctors

and patients, traditionally fighting to reduce uncertainty, must also learn to cope with possibly reaching the goal of getting certainty.

FUNDING AND FORMAL STANDARDS

In contrast with therapeutic research, and despite the clinical fact that accurate dia-prognosis is the key to effective management, the funding possibilities for dia-prognostic research are generally poor. Feinstein's plea for improvement is hitting the target. Such improvement would be stimulated by adopting formal standards for evaluation of dia-prognostic procedures, to control acceptance, maintenance, and substitution in the healthcare market. This also requires high quality and transparency of evaluation reports. The recent initiative taken by the Amsterdam Academic Medical Centre to reach international agreement on Standards for Reporting Diagnostic Accuracy (STARD) deserves therefore full support from the scientific and healthcare community.

J Epidemiol Community Health 2002;**56**:340–341

Author's affiliation
J A Knottnerus , Netherlands School of Primary Care Research, University of Maastricht, PO Box 616, 6200 MD Maastricht, Netherlands

Correspondence to: Professor J A Knottnerus; andre.knottnerus@hag.unimaas.nl

REFERENCES

1 **Feinstein AR**. *Clinical epidemiology: the architecture of clinical research*. Philadelphia: WB Saunders, 1985.
2 **Ransohoff DF**, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;**299**:926–30.
3 **Knottnerus JA**. The effects of disease verification and referral on the relationship between symptoms and diseases. *Med Decis Making* 1987;**7**:139–48.
4 **Begg CB**. Biases in the assessment of diagnostic tests. *Stat Med* 1987;**6**:411–23.
5 **Knottnerus JA**, Leffers P. The influence of referral patterns on the characteristics of diagnostic tests. *J Clin Epidemiol* 1992;**45**:1143–54.
6 **Moons KGM**, van Es GA, Deckers JW, *et al*. Limitations of sensitivity, specificity, likelihood ratio, and Bayes's theorem in assessing diagnostic probabilities: a clinical example. *Epidemiology* 1997;**8**:12 –17.
7 **Sheps SB**, Schechter MT. The assessment of diagnostic tests. A survey of current medical research. *JAMA* 1984;**252**:2318–22.
8 **Reid ML**, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic research. Getting better but still not good. *JAMA* 1995;**274**:645–51.
9 **Lijmer JG**, Mol BW, Heisterkamp S, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6.
10 **Sackett DL**, Haynes RB. The architecture of diagnostic research. In: Knottnerus JA, ed. *The evidence base of clinical diagnosis*. London: BMJ Books, 2002.
11 **Bogardus ST Jr**, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA* 1999;**281**:1919–26.