Evaluation of the research methodology in genetic, molecular and proteomic tests

Blanca Lumbreras / Inmaculada Jarrín / Ildefonso Hernández Aguado
Department of Public Health, History of Science and Gynaecology.
Universidad Miguel Hernández, Alicante, Spain.

**(Valoración de la metodología de la investigación
en pruebas de genética, molecular y proteómica)**

Abstract

Introduction: Advances in genomic analysis technologies have led to the development of new diagnostic tests with clinical application. Therefore, as in other diagnostic fields, awareness of the methodological limitations of genetic investigation will facilitate the application of the results.

Methods: 44 articles which studied the diagnostic accuracy of genetic, molecular and proteomic tests, and published in *JAMA*, *Lancet*, *New England Journal of Medicine, Cancer Research* y *Clinical Cancer Research* from 2002 to June 2005 were analysed. 24 methodological criteria of the STARD guide (*Standards for Reporting of Diagnostic Accuracy)* were applied.

Results: The mean number of methodological criteria satisfied was 9.8 (95%, CI 8.8-10.6), with the greatest deficiencies observed in the aspects related to the description of patient selection, 9 (20%), the treatment of indeterminate results, 5 (11%), and the determination of the technique's reproducibility, 6 (13%). In contrast, a high frequency was observed in the description of the reference standard, 39 (87%), and the method used, 28 (62%).

Discussion: The articles evaluated fail to fulfil many of the quality requirements laid out in the STARD proposal, with the methodological quality being lower than in other diagnostic fields. The aspects most in need of improvement are those related to the description of patient selection and the determination of reproducibility. Research and progress in new genetic-molecular technologies require improved fulfilment of the epidemiological and clinical standards which are already applied by other diagnostic fields.

Key words: Genetic tests. Methodology. Sensitivity and specificity.

Resumen

Introducción: El avance en las tecnologías del análisis genómico ha supuesto el desarrollo de nuevas pruebas diagnósticas con potencial aplicación clínica. Así como ha sucedido en otros campos del diagnóstico, conocer las limitaciones metodológicas de la investigación en genética facilitará la aplicación de sus resultados.

Métodos: Se analizaron 44 artículos publicados en *JAMA*, *Lancet*, *New England Journal of Medicine, Cancer Research* y *Clinical Cancer Research,* de 2002 a junio de 2005, que estudiaban la exactitud diagnóstica de pruebas genéticas, moleculares y de proteómica. Se aplicaron 24 criterios metodológicos de la guía STARD (Standards for Reporting of Diagnostic Accuracy).

Resultados: La media de cumplimiento de los criterios aplicados fue de 9,8 (intervalo de confianza [IC] del 95%, 8,8-10,6), y las mayores deficiencias se observaron en los aspectos relacionados con la descripción de la selección de los pacientes (9; 20%), el tratamiento de los resultados indeterminados (5; 11%) y la determinación de la reproducibilidad de la técnica (6; 13%). En cambio, sí se observó una alta frecuencia en la descripción del patrón de referencia (39; 87%) y del método usado (28; 62%).

Discusión: Los artículos evaluados incumplen gran parte de los requisitos de calidad recogidos en la propuesta STARD, observándose una peor calidad metodológica que en otros campos diagnósticos. Los aspectos que más necesitan mejorar son los relacionados con la descripción de la selección de los pacientes y la determinación de la reproducibilidad. La investigación y el avance en nuevas tecnologías genético-moleculares requieren un mejor cumplimiento de los estándares epidemiológicos y clínicos que ya se aplican a otros campos diagnósticos.

Palabras clave: Pruebas genéticas. Metodología. Sensibilidad y especificidad.

Correspondence: Dra. Blanca Lumbreras.
Departamento de Salud Pública, Historia de la Ciencia y Ginecología. Campus de San Juan. Facultad de Medicina.
Universidad Miguel Hernández.
Ctra. de Valencia, km 8,7. 03550 San Juan de Alicante.
España.
Correo electrónico: blumbreras@umh.es

Introduction

The advances derived from the Human Genome Project have led to the development of new kinds of diagnostic tests, genetic, molecular or proteomic, which can be incorporated into clinical practice[1-4]. In Spain currently there are tests based on DNA analysis for the diagnosis or prognosis of 214 genetic

diseases[5], and it is foreseen that in the coming years this number of tests will multiply. This has given rise to the elaboration of evaluation frameworks for the incorporation of new genetic tests to the list of services of health systems[6]. The appearance of these new tests creates expectations, which often are not corroborated in clinical practice. This is due to the fact that sometimes the results of basic research are publicised without awaiting confirmation from the results of clinical research; at other times the clinical validation of the tests lacks methodological rigour, precisely for their having been developed in basic research environments without the collaboration of investigators with experience in clinical and epidemiological research. In any case the new genetic-molecular technologies should not be introduced for a specific clinical task (screening and genetic counsel, diagnosis, risk evaluation, etc.), without prior examination of both their validity for this specific purpose as well as their effects in health terms[1].

A methodological deficit in the quality of research in other diagnostic areas has been described[7-12]. In genetic-molecular research, however, a single work analysed the methodology of the studies published in four international medical journals, showing that 63% of them fulfilled only one of the required methodological criteria[13].

In order to improve the articles on diagnostic research and, as a result, the scientific quality of this research, the STARD (Standards for Reporting of Diagnostic Accuracy)[14] was initiated, and its recommendations have been incorporated by most of the scientific journals. Moreover, Little and collaborators[15] presented a list of standards shared in part by STARD, but with some aimed solely at research into markers based on genotypes, which proposed to serve as a guide for authors as well as editors and reviewers of genetics articles.

The works dealing with genetic-molecular tests have a series of specificities, but these do not exclude the need for fulfilment of the quality criteria that are demanded in other diagnostic studies. Furthermore, the quality achieved in the studies can differ depending on whether the journal in which they are published belongs to the clinical or diagnostic field. To determine whether genetic-molecular research achieves the required quality, the studies on genetic-molecular diagnostic accuracy published since the year 2002 in four very important international clinical journals and in two journals of the American Association for Cancer Research (AACR), characterised by including research on genetic-molecular tests, were evaluated by applying the criteria of the STARD[14] guide.

Methods

All original articles studying the diagnostic accuracy of genetic, molecular and proteomic tests from Jan-

uary 2002 to June 2005 in four international clinical journals (*JAMA*, *British Medical Journal*, *Lancet* and *New England Journal of Medicine*) and 2 journals centred on the field of cancer research belonging to the American Association for Cancer Research (AACR) (*Cancer Research* and *Clinical Cancer Research*) were reviewed.

Selection of articles

A search was conducted through the Medline computerised bibliographic system using the search strategy employed by Devillé and collaborators[16], combining the MESH terms «sensitivity and specificity» with the words «false negative» and «accuracy». In order to improve the sensitivity of the search it was widened with the MESH term «area under the curve» and the words «diagnostic odds ratio» and «likelihood ratio». After reading the summaries, all the articles in which human genetic material was analysed at molecular level and all proteomic studies were selected, without limiting the search to any particular laboratory technique or clinical condition.

Methodological standards

24 of the 25 criteria belonging to the STARD[14] guide were applied (the first standard of this guide was omitted for being the criterion for inclusion in our study). The given definitions were followed in all of them. Although they can be consulted in detail in the reference publications, the following shows the methodological standards grouped into four blocks:

1. Introduction. Within the study aims, an estimate of the diagnostic accuracy of a test should be included, or the comparison of that accuracy among the tests or participating groups (1).

2. Methods. The inclusion and exclusion criteria of the participants in the study should appear (2) and the place where the data or samples were gathered, as well as specification of the reason for their choice (3). In the case of it not being a consecutive sample of patients then how they were assembled should be explained (4), as well as the directionality of the study (5). On the test being evaluated, the reference standard used should be defined (6) and technical specifications given, which include how and when the measurements were made, for both the diagnostic test and the reference standard (7). The units and/or categories of the results of the diagnostic test and the reference standard should also be described (8). Additionally, the number, training and experience of the people who have carried out and interpreted the tests and the reference standard should be specified (9), and if the results have undergone blind

reviewing (10). The statistical methods used to calculate and compare measurements of diagnostic accuracy and to quantify the standard error should be described (11), as too should the methods for calculating the reproducibility of the tests (12).

3. Results. This section compiles characteristics of when the study was carried out (13), the clinical and demographic characteristics of the study population (14) and the number of participants which satisfied the selection criteria, and which then received or not the diagnostic test and/or the reference standard (this can be illustrated with a flow diagram) (15). Also included is the time interval between the application of the diagnostic test, the reference standard and any other treatment administered between these (16), the distribution of the severity of the disease among the patients (17), a table with the results of the diagnostic tests in comparison with those obtained from the reference standard (18) and the communication of any adverse effect produced during the study (19). Lastly, the diagnostic accuracy and the statistical precision should be estimated (20), how indeterminate results have been treated (21), the estimation of the variability of the diagnostic accuracy for each subgroup of participants, researchers or centres (22), and the calculation of the reproducibility of the technique (23).

4. Discussion. In this section it is important to comment on the clinical applicability of the study findings (24).

Observer variability

In order to calculate the inter-observer variability in the application of the criteria, 4 of the 44 works were selected at random and reviewed independently by the two observers. The degree of agreement reached between them was 86% (confidence interval [CI] 95%, 78-92).

The creation and management of the data base and the statistical analyses of the results were performed with the SPSS statistical package (version 12 for Windows Inc., Chicago, Illinois).

Results

In the study period 44 articles fulfilled the inclusion criteria: 12 from *Lancet*, 6 from *JAMA*, 2 from *New England Journal of Medicine*, none from *BMJ*, 14 from *Cancer Research* and 10 from *Clinical Cancer Research*.

All of the articles reviewed evaluated tests centred on the study of cancer, with prostate cancer, 10 (22.8%), and breast cancer, 7 (15.9%), being the most frequent (table 1). Most of the diagnostic tests evaluated, 37 (84%), were genetic, both microarray systems,

15 (41%), as well as genetic sequencing analysis, 12 (59%), and only 7 (16%) were proteomic. Of all the works, 7 (16%), were prognostic in nature and most, 37 (84%), were diagnostic.

Mean fulfilment of the 24 methodological standards applied was 9.8 (CI 95%, 8.8-10.6), and none of the 44 articles evaluated satisfied more than 18 criteria. No statistically significant differences were observed in the fulfilment of the criteria before and after the year 2003 (publication of the STARD guide[14]).

Analysing the criteria individually, the greatest deficiencies were found in the material and methods section (table 2). Only 11 works (24%) specified the determination of this diagnostic accuracy as being among their objectives. Also low was the frequency with which inclusion and exclusion criteria and the place of origin of the subjects was specified (23%); the way in which these were selected (20%), or whether this had been planned before or after the performance of the diagnostic test (27%). Regarding the performance of the test, there was a low level of fulfilment in evaluators description, 1 work (2%), and whether the interpretation of the results was blinded, 12 articles (27%). The description of the methods for studying the reproducibility of the test was only communicated in 10 studies (23%) and the results obtained being shown in only 6 articles (13%).

With regard to the presentation of the results (table 3), most defined the clinical and demographic characteristics of the study population, 34 (77%), but the reasons for which some patients who fulfilled the inclusion criteria but then did not continue in the study are hardly given, 11 (25%). Only 12 studies (27%) investigated whether the diagnostic sensitivity and specificity varied according to relevant clinical subgroups. Communication and analysis of the indeterminate or imprecise results was very infrequent, 5 (11%).

If we compare the fulfilment of the criteria among the clinical journals and those dedicated to cancer re-

Table 1. Description of the diseases evaluated in the works on investigation of genetic, molecular and proteomic tests analysed in *JAMA, New England Journal of Medicine, Lancet, Cancer Research* and *Clinical Cancer Research* (2002-2005)

Diseases	n	%
1. Prostate cancer	10	22.8
2. Breast cancer	7	15.9
3. Renal cancer	5	11.4
4. Ovary cancer	4	9.1
5. Hepatocellular carcinoma	4	9.1
6. Melanoma	4	9.1
7. Cancer of the colon	4	9.1
8. Lung cancer	2	4.5
9. Neuroblastoma	2	4.5
10. Bladder cancer	2	4.5

Table 2. Fulfilment of the criteria relative to «Material and methods» for articles on investigation of genetic, molecular and proteomic tests in *JAMA, New England Journal of Medicine, Lancet, Cancer Research* and *Clinical Cancer Research* (2002-2005)

Criteria	*JAMA, Lancet* and *New England Journal of Medicine* (n = 20)	*Cancer Research* and *Clinical Cancer Research* (n = 24)	Total (n = 44)
1. Objectives	5 (25%)	6 (25%)	11 (24%)
Participants			
2. Inclusion and exclusion criteria	8 (40%)	2 (8%)	10 (23%)
3. Assembling patients	18 (90%)	22 (92%)	40 (89%)
4. Sample selection	6 (30%)	3 (13%)	9 (20%)
5. Direction of the study	7 (35%)	5 (21%)	12 (27%)
Test			
6. Description of the reference pattern	17 (85%)	22 (92%)	39 (87%)
7. Technical specifications: how and when	7 (35%)[a]	21 (88%)[a]	28 (62%)
8. Definition of the units	7 (35%) [a]	20 (83%)[a]	27 (61%)
9. People carrying out the test	1 (5%)	0	1 (2%)
10. Blind review	11 (55%)[a]	1 (4%)[a]	12 (27%)
Statistical methods			
11. Statistical methods	9 (45%)	17 (71%)	26 (58%)
12. Reproducibility methods	4 (20%)	6 (25%)	10 (23%)

[a]Statistically significant differences ($p < 0.05$) between the two groups of journals calculated by the Fisher test.

search (tables 2 and 3), it can be seen that statistically significant differences exist in certain standards. The cancer-related journals fulfilled the definition of the technical specifications of the test better than the clinical publications, 21 (88%) and 7 (35%) respectively, and the specification of the units employed in the test, 20 (83%) compared to 7 (35%). They also expressed the correspondence of the results of the reference standard with the evaluated test more frequently, 22 (92%) compared to 8 (40%), and the study of the variation of accuracy according to relevant clinical subgroups, 10 (42%) compared to 2 (10%). However, the clinical journals showed greater fulfilment of the standards related to the blind reviewing of the results compared to the cancer journals, 11 (55%) and 1 (4%) respectively, and to the presentation of the estimates of accuracy and their statistical precision, 12 (60%) compared to 6 (25%).

Discussion

The articles on genetic-molecular diagnostic tests published recently (2002-2005) in the best international journals of clinical medicine and in two important publications of cancer research, fail to satisfy most of the quality requirements assembled in the STARD[14] proposal. This shortcoming could be a reflection of severe limitations in the research methodology, or carelessness in the preparation of the articles and their editorial process, or both. If we compare with other diagnostic fields[9-12], it can be seen that genetic, proteomic and molecular studies show inferior methodological quality, while for example in the laboratory field, and due fundamentally to the efforts of different authors and editorial groups, there has been a gradual improvement in their quality[11].

The methodological quality found in both the clinical journals as well as those dedicated to cancer research is similar, although with some specificity. The most analytical aspects, like the specification of the technical characteristics, are given more attention to in the cancer journals; while others more related to the way in which to carry out the study, as for example the blind reviewing of the results obtained, are better fulfilled in the clinical journals.

The only antecedent in the field of genetic-molecular diagnosis of this kind of investigation was a review carried out in 1995 on the same four clinical journals as those analysed in this work[13]. Although the same methodological criteria were not used, those employed at that time are included in the STARD guide, and are therefore comparable. As in this review, no work fulfilled an acceptable number of methodological standards, the calculation of the genetic test's reproducibility was unusual and there was a general lack of blind reviewing in the interpretation of the results. Hence it would appear that genetic-molecular diagnostic studies are more impermeable to clinical-epidemiological advances.

Our intention of facilitating comparison with previous works led us to the application of the same restrictive selection criteria as employed previously[7,9-12], and as a result to a reduced sample of articles, a fact which li-

371

Table 3. Fulfilment of the criteria relative to the results and discussion of the articles on genetic, molecular and proteomic diagnostic tests in _JAMA, New England Journal of Medicine, Lancet, Cancer Research_ and _Clinical Cancer Research_ (2002-2005)

Criteria	_JAMA, Lancet_ and _New England Journal of Medicine_ (n = 20)	_Cancer Research_ and _Clinical Cancer Research_ (n = 24)	Total (n = 44)
Participants			
13. Study dates	12 (60%)	6 (25%)	18 (41%)
14. Description of the subjects' characteristics	15 (75%)	19 (79%)	34 (77%)
15. Presentation of the flow diagram	8 (40%)	3 (13%)	11 (25%)
Results of the test			
16. Time between the test and the reference standard	4 (20%)	11 (46%)	15 (34%)
17. Distribution of the disease	16 (80%)	21 (86%)	37 (84%)
18. Results of the test and reference pattern	8 (40%)[a]	22 (92%)[a]	30 (68%)
19. Communication of adverse effects	0	0	0
Estimates			
20. Presentation of the accuracy estimates and their statistical precision	12 (60%)[a]	6 (25%)[a]	18 (41%)
21. Communication of the indeterminate results	3 (215%)	2 (8%)	5 (11%)
22. Studies of the variation of the accuracy according to relevant clinical subgroups	2 (10%)[a]	10 (42%)[a]	12 (27%)
23. Determination of the reproducibility	3 (15%)	3 (13%)	6 (13%)
Discussion			
24. Evaluation of the clinical applicability	9 (45%)	9 (42%)	18 (42%)

[a]Statistically significant differences ($p < 0.05$) calculated for χ^2 between the two groups of journals calculated by the Fisher test.

mits the reach of the results. Nevertheless, by including all the original articles which have been published on the field of genetic and molecular diagnosis in the best international journals of clinical medicine and in the two publications of the American Association of Cancer Research (for containing an important part of the works published on the same matter), an adequate perspective of the quality of research and publication of this kind of study is offered. The guide employed for the evaluation of the articles (STARD) is very recent but has been rapidly disseminated among most journals which publish diagnostic evaluations; in fact, two of them (_JAMA_ and _Lancet_) have included it among the rules for authors[17,18], although not so _New England Journal of Medicine_ or the two cancer journals. It could be argued that there is little experience in its application, and it is true that there may be doubts about the applicability of some of its criteria due to the lack of precise indications. However, available experience shows that concordance on its application is high among observers[11,12], as it has been too in the sample carried out in the present study. Moreover, many of the STARD criteria have already been applied in previous methodological guides[7-12], which has facilitated its use. Studies on diagnosis in genomics, proteomics and related fields have some specificities in their validation[19] which the STARD guide does not include, and although various initiatives have originated for the elaboration of a set of guidelines suitable for the validation of these tests[15], no definitive one has been drawn

up yet. This fact represents a limitation in that some intrinsic problems have not been evaluated, although they are not exclusive to this field of research. One example is «over-fitting», which occurs for example when a proteomics profile is proposed and selected from a broader group after countless discrimination trials, which affects external validity. This problem is controlled by determining the reproducibility of the discriminatory pattern observed in samples independent of the original. If the discrimination pattern cannot be reproduced in other samples, the existence of «overfitting» is probable[20].

Another limitation to bear in mind is the study period, since the STARD criteria were first published early in the year 2003[14], and this study recompiles articles from 2002 to 2005. Certainly it is early for the initiative to have had effect, at least in the journals which have included it in the instructions for their authors; and our study does not have the statistical capacity to establish comparisons. However, the fact that it was not published does not explain the low quality observed, as many of its criteria are essential requirements in most observational studies already partially compiled —the most relevant ones— in other guides and previous publications.

Not all the requirements have the same importance in terms of their effects on the validity of the evaluated studies. Bearing in mind the main characteristics that a genetic test should possess for its correct use[21], it is important to underline the key aspects of validity which

have shown themselves to be the most deficient. In first place are the analytical characteristics of the test evaluated, as due to the current rapid development of genetic tests for the same test there may be different techniques and one should know which of them shows the most acceptable values of diagnostic accuracy. Hence, the scanty calculation of the technique's reproducibility in the investigations evaluated is notable. In previous reviews this concept was already considered essential[7-12], and yet in the field of genetics or proteomics, where the value of the determinations will also depend on the capacity to give the same result when applied to the same patients under the same conditions[19], it has been neglected all too often.

Secondly, another key methodological deficiency, and one which limits the applicability of the results, is the inadequate description of the inclusion and exclusion criteria of the subjects participating in the study, these barely appearing in less than half of the works evaluated. This lack of information impedes detailed consideration of the clinical applicability of the test evaluated, which is an indispensable element in this kind of investigation[21].

To summarize, the future projection of the new genetic-molecular technologies in the field of diagnosis and precocious detection is very limited by the lack of rigour in the research on these tests. These problems are at the root of the reiterated false expectations which are created around these diagnostic tests. If an effective interdisciplinary relationship (basic, clinical, epidemiological) which facilitated the rigorous development of these genetic-molecular diagnostic tests existed then a lot of unnecessary research would be avoided. Furthermore, adherence to the methodological standards would allow the genetic-molecular and proteomic tests to become very useful tools in clinical and health service research, especially in the field of cancer[22].

Acknowledgements

References

1. Schwartz MK. Genetic testing and the clinical laboratory improvement amendments of 1988: present and future. Clin Chem. 1999;45:739-45.
2. Higashi MK, Veenstra DL. Managed care in the genomics era: assessing the cost effectiveness of genetic tests. Am J Manag Care. 2003;9:493-500.
3. Milunsky A. Commercialization of clinical genetic laboratory services: in whose best interest? Obstet Gynecol. 1993;81:627-9.
4. Korf BR. Advances in molecular diagnosis. Curr Opin Obstet Gynecol. 1996;8:130-4.
5. Rueda J, Briones E. Servicios de diagnóstico genético para las enfermedades hereditarias en España. EUR 20516 EN. Institute for Prospective Technological Studies, 2002.
6. Márquez S, Briones E. Marco para la evaluación de las pruebas genéticas en el Sistema Sanitario Público de Andalucía. Informe 2/2005.
7. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. JAMA. 1995;274:645-51.
8. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, Van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282:1061-6.
9. Ramos Rincón JM, Hernández Aguado I. Investigación en pruebas diagnósticas en Medicina Clínica. Una evaluación de la metodología. Med Clin (Barc). 1998;111:129-34.
10. Ramos JM, Hernández I. Métodos para evaluar pruebas diagnósticas en enfermedades infecciosas y microbiología clínica. Enferm Infecc Microbiol Clin. 1998;16:179-84.
11. Lumbreras-Lacarra B, Ramos-Rincón JM, Hernández-Aguado I. Methodology in diagnostic laboratory test research in clinical chemistry and clinical chemistry and laboratory medicine. Clin Chem. 2004;50:530-6.
12. Lumbreras Lacarra B, Ramos Rincón JM, Hernández Aguado I. Evaluación de la investigación metodológica en pruebas diagnósticas de laboratorio en Revista Clínica Española y Medicina Clínica. Rev Clin Esp. 2004;204:472-6.
13. Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. JAMA. 1999;281:1919-26.
14. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. Clin Chem. 2003;49:1-6.
15. Little J, Bradley L, Bray MS, Clyne M, Dorman J, Ellsworth DL, et al. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol. 2002;156:300-10.
16. Deville WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. J Clin Epidemiol 2000;53:65-9.
17. Rennie D. Improving reports of studies of diagnostic tests: the STARD initiative. JAMA. 2003;289:89-90.
18. Bossuyt PM, Reitsma JB. Standards for Reporting of Diagnostic Accuracy. The STARD initiative. Lancet. 2003;361:71.
19. Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Cancer. 2004;4:309-14.
20. Ransohoff DF. Lessons from controversy: ovarian cancer screening and serum proteomics. J Natl Cancer Ins.t 2005; 97:315-9.
21. Burke W. Genetic testing. N Engl J Med 2002;347:1867-75.
22. Porta M, Fernández E, Alguacil J. Semiology, proteomics and the early detection of symptomatic cancer. J Clin Epidemiol. 2003;56:815-9.