

Evaluation of formulae for CSF IgG synthesis using data obtained from two methods: importance of receiver operator characteristic curve analysis

S A McMillan, J P Douglas, A G Droogan, S A Hawkins

Abstract

Aims—To determine the clinical performance of three cerebrospinal fluid (CSF) IgG synthesis formulae using data obtained from two quantitation methods.

Methods—Receiver operator characteristic (ROC) analysis and decision index plots were used to compare a rate nephelometric (RN) and a rocket immunoelectrophoretic (RIEP) method for quantitating albumin and IgG for use in CSF IgG synthesis formulae. Further analysis was used to determine the most clinically accurate of these formulae for a diagnosis of multiple sclerosis with regard to technical accuracy and cost effectiveness.

Results—Values for albumin and IgG determined by RN gave better sensitivities and specificities than the RIEP method when applied to all three formulae; however, when the 95% confidence limits were considered, the difference was not significant. Using the RN method with an agreed "rule in" threshold value of 90% specificity, the IgG index gave the best clinical performance.

Conclusion—ROC curve analysis and decision index plots provide valuable tools in assessing and comparing the clinical performance of new and existing laboratory assays.

(*J Clin Pathol* 1996;49:24-28)

Keywords: receiver operator characteristics, multiple sclerosis, intrathecal IgG synthesis.

Multiple sclerosis is a relapsing and remitting demyelinating disease of the central nervous system in which early diagnosis is helpful in its clinical management. Although no specific laboratory tests exist, there are a number of biochemical tests which can be used as an aid to support its diagnosis. These tests demonstrate the intrathecal synthesis of IgG and include the detection of oligoclonal IgG bands in the cerebrospinal fluid (CSF)¹ and the use of various formulae to indicate CSF IgG synthesis.² The latter include numerous different formulae, each varying in sensitivity and specificity,² which use data obtained from the quantitation of albumin and IgG in patients' CSF and serum. Various methods have been used to quantitate IgG and albumin in paired serum

and CSF samples including electroimmunodiffusion,³ radial immunodiffusion,⁴ rocket immunoelectrophoresis,⁵ and automated nephelometric immunoassay.⁶ The presence of oligoclonal bands seems to be the most informative of these laboratory tests; however, many diagnostic laboratories use a combination of oligoclonal band detection and one or more of the formulae.

A diagnostic laboratory should always be reviewing the clinical performance and cost effectiveness of existing assays in the light of newer and better ones which may be added to the laboratory's repertoire of tests and the old tests eliminated. A laboratory test's clinical performance is its ability to classify subjects correctly into clinically relevant groups.⁷ The diagnostic accuracy of a test is one way of measuring this. Receiver operating characteristic (ROC) curves are an excellent method of assessing diagnostic accuracy of one or a number of tests by demonstrating the limits of a test's ability to discriminate between different health states over a range of test result values. Once a ROC curve has been plotted, other decision analyses can be made, including the appropriate decision levels to be used.^{7,8}

Here we describe the use of ROC curve analysis in evaluating an automated method for the quantitation of IgG and albumin (rate nephelometry). Other laboratory characteristics of method performance were also used to help in the decision process. Further analysis was used to re-evaluate the clinical performance of the intrathecal IgG synthesis formula used in our laboratory and to determine the best decision values for the group of patients being studied.

Methods

SAMPLE/PATIENT POPULATION

Paired serum and CSF samples from 175 consecutive patients with suspected multiple sclerosis were sent from Neurology outpatient clinics to the Regional Immunology Laboratory for routine intrathecal IgG synthesis analysis and oligoclonal band detection over a period of six months. Laboratory analysis included the quantitation of albumin and IgG, the calculation of a CSF IgG/albumin ratio (IgG ratio),⁹ the CSF IgG index,¹⁰ the CSF IgG synthesis rate,¹¹ and the detection of oligoclonal IgG bands in the CSF.¹² The clinical diagnosis of these patients was determined by one of us

Regional Immunology
Laboratory, Belfast
City Hospital, Belfast
BT9 7AD
S A McMillan
J P Douglas

Department of
Neurology, Royal
Victoria Hospital,
Belfast BT12 6BA
A G Droogan
S A Hawkins

Correspondence to:
Dr S A McMillan.

Accepted for publication
5 September 1995

(AGD) by personal clinical contact and chart review at least 12 months after samples were received in the laboratory. Seventy nine patients were classified as definite multiple sclerosis according to the criteria of Poser *et al.*,¹³ 36 as inflammatory neurological disease and 60 as non-inflammatory neurological disease.

CALCULATION OF CSF IgG SYNTHESIS FORMULAE

Rocket immunoelectrophoresis following the method of Laurell⁵ and rate nephelometry using a Beckman Array Protein system with Beckman standards and controls⁶ were used to quantitate IgG and albumin in paired serum and CSF samples. The CSF IgG ratio was calculated as the ratio of the concentration of IgG to albumin in the CSF.⁹

The CSF IgG index was calculated as follows¹⁰:

$$\text{IgG Index} = \frac{\text{Albumin (serum)}}{\text{IgG (serum)}} \times \frac{\text{IgG (CSF)}}{\text{albumin (CSF)}}$$

The IgG synthesis rate was calculated using the following formula,¹¹ and the result expressed in mg.

$$\begin{aligned} \text{IgG/day:} \\ \left[\text{IgG CSF} - \frac{\text{IgG serum}}{369} - \left(\text{Alb CSF} - \frac{\text{Alb serum}}{230} \right) \right. \\ \left. \times \frac{\text{IgG serum (0.43)}}{\text{Alb serum}} \right] \times 5 \end{aligned}$$

where Alb = albumin.

COMPARISON OF METHODS

ROC curves were constructed for each of the three formulae using results from the two methods for quantitating albumin and IgG (rate nephelometry and rocket immunoelectrophoresis). Results were divided into either positive or negative at a range of cut off values for the IgG index, IgG ratio and IgG synthesis rate when a clinical diagnosis of multiple sclerosis was used as the gold standard. ROC curves are constructed by plotting the sensitivity (true positive rate) against 1 – specificity (false positive rate) at a range of cut off values,⁷ giving a visual comparison of the different methods. The area under the ROC curve was also used to quantify the diagnostic accuracy of different formulae as it expresses the performance as a single number.⁷ This value was determined by weighing each plot and calculating this value (including standard error) as a ratio of the weight of the total area of the graph.

Decision index plots were also determined for each formula using the two methods of quantitation to obtain the optimum threshold value. This index gives the best compromise between the true positive rate (sensitivity) and false positive rate (1 – specificity) and is calculated as sensitivity + specificity – 1 (Youden J index) for each cut off value.¹⁴ The sensitivity and specificity for each formula using the two

methods with their respective 95% confidence limits, obtained from Geigy Scientific tables,¹⁵ was determined at each optimum threshold value.

DETERMINATION OF THE BEST DECISION LEVEL

A decision threshold of 90% specificity was determined for the three formulae using quantitation by rate nephelometry. Sensitivity and positive and negative predictive values and test efficiency for the three formulae were also determined at this decision level (cut off level).¹⁴

DEFINITIONS

Sensitivity is defined as the proportion of true positive test results in the patients with a diagnosis of multiple sclerosis (also known as the true positive rate (TPR)) and the specificity is the percentage of true negative test results in the group of patients without multiple sclerosis. The false positive rate (FPR) is 1 – specificity. The positive and negative predictive values are defined as the proportion of all positive or negative tests results which correctly indicate multiple sclerosis or not multiple sclerosis, respectively. The positive (PVP) and negative (PVN) predictive values were calculated using the following formulae:

$$\text{PVP} = \frac{(\text{Prevalence})(\text{TPR})}{[(\text{Prevalence})(\text{TPR}) + \text{FPR}(1 - \text{Prevalence})]}$$

$$\text{PVN} = \frac{(1 - \text{Prevalence})(1 - \text{FPR})}{[(1 - \text{Prevalence})(1 - \text{FPR}) + \text{Prevalence}(1 - \text{TPR})]}$$

The efficiency of the test is defined as the proportion of all results (both positive and negative) that are correct and is calculated as follows:

$$E = \frac{\text{No. of true positives} + \text{No. of true negatives}}{\text{All patients in study (disease and non-disease)}} \times 100$$

where E = efficiency.

LABORATORY CHARACTERISTICS OF RATE NEPHELOMETRY AND ROCKET IMMUNOELECTROPHORESIS

Certain laboratory characteristics were also considered in assessing the best method. These included assay cost, precision and reproducibility, hands on time, time to result, standardisation, quality control, and technical difficulty. Direct and indirect costs were used to calculate the price of tests. Direct cost relating to staff grade required to carry out the assays and the indirect costs of supervisory staff, computer facilities, research and development, and laboratory and hospital overheads were common to both methods and therefore were not included. Only the cost of consumables and of the service contracts/breakdown charges was used. Precision and reproducibility were determined by calculation of the intra- and interbatch mean, standard deviation and coefficient of variation for each method. Hands on time was the length of time one person spent in performing each method and the time to result was the length of time from the start

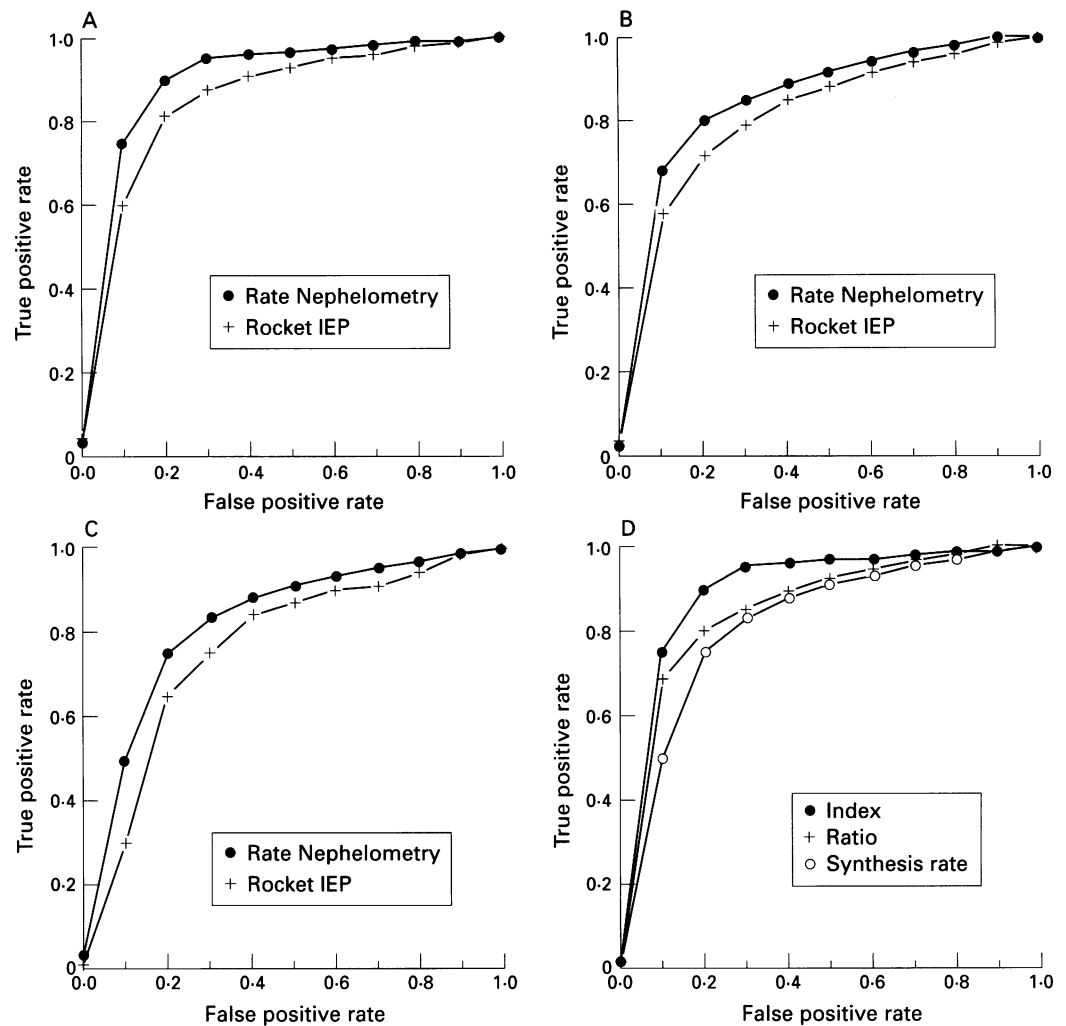


Figure 1 ROC curves of the rate nephelometric and rocket immunoelectrophoretic (Rocket IEP) quantitation methods for the IgG index, IgG ratio and IgG synthesis rate and the combined three formulae using rate nephelometry. (A) IgG index; (B) IgG ratio; (C) IgG synthesis rate; (D) IgG synthesis formulae.

of the assay to reporting of a result. These times were for batches of equal sample numbers. Subjective analysis of the ease of assay standardisation and of obtaining adequate quality control together with experimental (technical) difficulty in performing the methods was also taken into account.

Results

COMPARISON OF METHODS

The ROC curves for the three IgG synthesis formulae using results obtained by rate nephelometry and rocket immunoelectrophoresis are shown in fig 1. Both the IgG index and IgG ratio by the two methods showed moderate to high test accuracy as demonstrated by the closeness of the curve to the left and top margins. For each of the formulae, the ROC curves obtained using results generated by rate nephelometry gave better sensitivities and speci-

ficities over a range of cut off values than those obtained by rocket immunoelectrophoresis. The area under the ROC curves again showed that rate nephelometry gave higher values, indicating that this method was better; however, taking into account the standard error there was slight overlap in values (table 1).

Comparison of the decision plots again showed that the formulae using rate nephelometry gave higher decision index values than with rocket immunoelectrophoresis (fig 2; only IgG index and IgG ratio shown). However, the optimum decision value for each of the formulae was different using the two methods; the best value for the IgG index using rate nephelometry was 0.59 compared with 0.81 for rocket immunoelectrophoresis and for the IgG ratio these values were 0.19 and 0.26, respectively. The sensitivity and specificity at the optimum decision value of each formula using the two methods demonstrated that rate nephelometry always gave higher sensitivities with somewhat similar specificities. When the 95% confidence limits were applied there was no clear cut difference between sensitivities and specificities using the two methods (table 2). Using these various analyses it can be seen that the rate nephelometry method gave a slightly

Table 1 Comparison of area under curve (standard error) for IgG synthesis formulae using values obtained by rate nephelometry and rocket immunoelectrophoresis

Formula	Rate nephelometry		Rocket immunoelectrophoresis	
Index	0.902	(0.033)	0.857	(0.031)
Ratio	0.860	(0.031)	0.830	(0.029)
IgG synthesis rate	0.817	(0.031)	0.758	(0.032)

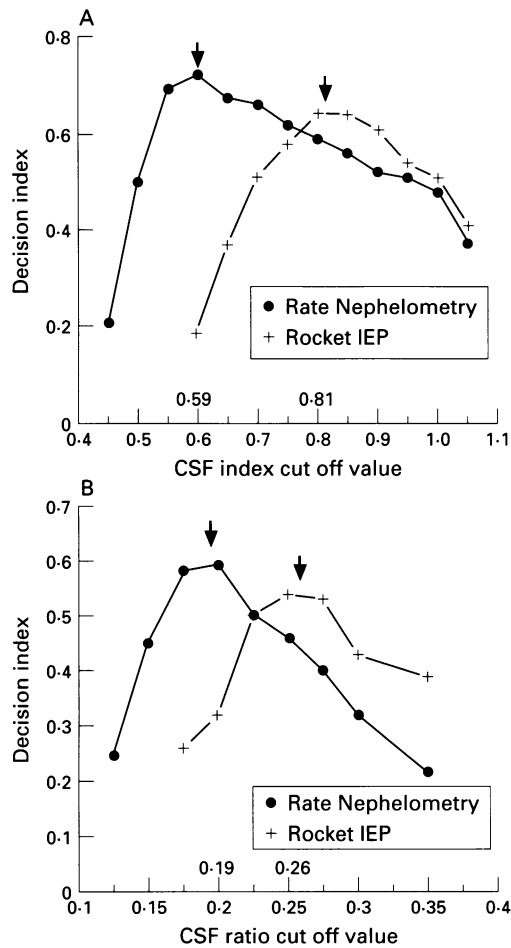


Figure 2 Decision index plots for (A) IgG index and (B) IgG ratio using rate nephelometry and rocket immunoelectrophoresis (Rocket IEP). Optimum cut off values for each formula are shown by arrows.

better clinical performance than rocket immunoelectrophoresis for all three IgG synthesis formulae.

DETERMINATION OF OPTIMUM CUT OFF VALUES

In determining the best cut off value we decided to use a value for each of the three formulae which gave 90% specificity. The cut off values using rate nephelometry for the three formulae together with their sensitivities, positive and negative predictive values and test efficiencies

can be seen in table 3. It is clear from the results that the IgG index gave the best clinical performance demonstrated by higher positive and negative predictive values and test efficiency.

COMPARISON OF LABORATORY CHARACTERISTICS

The cost of a ratio/index was more expensive using rate nephelometry than rocket immunoelectrophoresis in terms of consumables used and equipment service/breakdown. Rate nephelometry showed greater precision and reproducibility than rocket immunoelectrophoresis; it also required less hands on time to perform and a result was obtained in one third less time. Both methods required a similar level of competency and responsibility from the operator. Standardisation and quality control for rocket immunoelectrophoresis was sensitive to operator variations whereas the automated rate nephelometry method had built in standardisation and quality control modules. Technically, the rocket immunoelectrophoresis method was more difficult to perform being subject to both experimental and operator variation (table 4).

Discussion

In diagnostic laboratory medicine there is always an on-going process of evaluating current assays in the light of newer ones. Here we describe the process we used to compare the clinical performance of two methods for quantitating IgG and albumin for use in three IgG synthesis formulae used as an aid to the diagnosis of multiple sclerosis. The process was first to calculate the sensitivity and specificity of each method for each formula using a clinical diagnosis of multiple sclerosis as our gold standard over a range of cut off values. From this data ROC curves were constructed and compared visually and quantitatively by the area under each curve. Decision index plots were also constructed and visualised and the optimum decision index (that cut off value giving the optimum value between true positive rate and false positive rate) was used to compare sensitivities and specificities at these values. It is worthwhile mentioning that the optimum de-

Table 2 Sensitivity and specificity with 95% confidence limits of the best threshold level for the three formulae using data obtained by rate nephelometry and rocket immunoelectrophoresis

Formula	Rate nephelometry			Rocket immunoelectrophoresis		
	Cut off value	Sensitivity	Specificity	Cut off value	Sensitivity	Specificity
Index	0.59	89%(79-95)	83%(74-90)	0.81	78%(67-86)	83%(74-90)
Ratio	0.19	79%(69-88)	81%(72-88)	0.26	75%(63-83)	82%(73-89)
IgG synthesis rate	4	78%(68-87)	79%(70-87)	7	70%(58-79)	82%(73-89)

Table 3 Cut off values giving a specificity of 90% with the three formulae using rate nephelometry

Formula	Specificity	Sensitivity	Cut off value	Predictive value		Efficiency
				Positive	Negative	
Index	90%	78%	0.7	86	83	96
Ratio	90%	62%	0.22	84	74	87
IgG synthesis rate	90%	47%	15	79	67	78

Table 4 Comparison of laboratory characteristics of the rate nephelometry and rocket immunoelectrophoresis

	Rate nephelometry	Rocket immunoelectrophoresis
Cost of consumables/assay	£3.46	£2.13
Service/breakdown cost/year	£91.60	£80.00
Precision/reproducibility	Good	Variable
Hands on time	45 minutes	130 minutes
Time to result	75 minutes	210 minutes
Standardisation/quality control	Objective	Subjective
Experimental difficulty	Easy	Reasonably difficult

cision values for the two methods were considerably different.

Our analysis suggests that the quantitation of IgG and albumin by rate nephelometry gave a slightly better clinical performance than rocket immunoelectrophoresis; however, when the 95% confidence limits were applied there was some overlap in the sensitivities and specificities of the two methods, indicating that confidence limits must be applied when comparing this sort of data.^{7,8}

Comparison of the two methods using other laboratory characteristics showed the rate nephelometry method to be more expensive in terms of the cost of consumables and servicing. However, the rate nephelometry method has the benefits and advantages of an automated technique with increased precision and reproducibility. The shorter technical time needed to perform the assay releases staff for other duties and the quicker time to result improves turnaround time. These and the ease of standardisation and improved quality control together with its experimental simplicity, in our view, far outweigh the extra direct costs, and indicate that rate nephelometry is a more technically accurate and cost effective method for the determination of a CSF IgG ratio/index.

Although ROC curve analysis is useful in establishing the clinical performance of a test, it does not indicate which is the best cut off (decision) level to use nor which sensitivity or specificity would be most appropriate for the patients studied. In most biochemical laboratory tests there is an overlap of results from patients with and without a certain condition. Due to this overlap it is not possible to choose a cut off value for a test which will give a 100% true positive rate and 0% true negative rate. As the data obtained from the three formulae give a continuous range of values, which is the best cut off value to use? Diagnostic tests can be used for three main purposes: discovery (screening), confirmation of a disease or exclusion of a disease. With each of these the optimum cut off value differs. A discovery or exclusion test requires a high sensitivity whereas a confirmatory test requires a high specificity. The CSF IgG synthesis tests requested by neurologists are used to help confirm the diagnosis of multiple sclerosis and therefore require a high specificity. In this evaluation we decided a specificity of 90% was required. This gave a cut off value of 0.70, 0.22 and 15mg/day for

the IgG index, IgG ratio and IgG synthesis rate, respectively. Comparison of the three formulae showed that at this specificity the sensitivity, predictive values and efficiency of the IgG synthesis rate was much lower than the other two and therefore we decided not to use this formula as part of the diagnostic repertoire.

In our study group the patients were pre-selected on clinical grounds towards those with a high suspicion of having multiple sclerosis. As the sensitivity and specificity do not directly indicate how a test will perform in a population it is essential to compare data when the prevalence of the disease in the study group is known. The positive and negative predictive values (which incorporate disease prevalence) can be used to give an indication of how well the test is performing. It is important, once the decision level has been established, to use predictive values to compare test performance. In our patient group use of the IgG index gave much better positive and negative predictive values and overall test efficiency than the other formulae at the 90% specificity level.

In conclusion, it is apparent from this study that correct evaluation analysis must be applied to any new assay being tested for diagnostic use.

- 1 Mehta PD. Diagnostic usefulness of cerebrospinal fluid in Multiple Sclerosis. *Crit Rev Clin Lab Sci* 1991;28:233-51.
- 2 Souverein JHM, Serree HMP, Peet R, Grenzebach Smit W, Bruyn GW. Intrathecal immunoglobulin synthesis. Comparison of various formulae with the "gold standard" of isoelectric focusing. *J Neurol Sci* 1991;102:11-16.
- 3 Tourtellotte WW, Tavolato B, Parker JA, Comiso P. Cerebrospinal fluid electroimmunodiffusion. An easy rapid sensitive reliable and valid method for the simultaneous determination of immunoglobulin G and albumin. *Arch Neurol* 1971;25:345-51.
- 4 Mancini G, Carbonara AO, Heremans JF. Immunochemical quantitation of antigens by single radial immunodiffusion. *Immunochemistry* 1965;2:235-54.
- 5 Laurell CB. Electroimmuno assay. *Scand J Clin Lab Invest* 1972;29(Suppl 124):21-37.
- 6 Gaudie J. IgG and albumin measurement in the evaluation of neurological disorders. *Beckman Bull* 1980;ICS-13.
- 7 Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;39:561-77.
- 8 Henderson AR. Assessing test accuracy and its clinical consequences: a primer for receiver operating characteristic curve analysis. *Ann Clin Biochem* 1993;30:521-39.
- 9 Tibbing G, Link H, Ohman S. Principles of albumin and IgG analysis in neurological disorders. I. Establishment of reference values. *Scand J Clin Lab Invest* 1977;37:385-90.
- 10 Delpach B, Lichtblau E. Etude quantitative des Immunoglobulines G et de l'albumine du liquide céphalo-rachidien. *Clin Chem Acta* 1972;37:15-23.
- 11 Tourtellotte WW. The cerebrospinal fluid in multiple sclerosis. In: Vinken PJ, Bruyn GW, eds. *Handbook of clinical neurology*. Vol 3. Amsterdam: Elsevier, 1985:79.
- 12 Walker RWH, Keir G, Johnston MH, Thompson EJ. A rapid method for detecting oligoclonal IgG in unconcentrated CSF, by agarose isoelectric focusing, transfer to cellulose nitrate and immunoperoxidase staining. *J Neuroimmunol* 1983;4:141-8.
- 13 Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for Multiple Sclerosis: guidelines for research protocols. *Ann Neurol* 1983;13:227-31.
- 14 Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32-5.
- 15 Diem K, Seldrup J. *Introduction to statistics, statistical tables and mathematical formulae—the binomial distribution*. Geigy Scientific Tables. Volume 2. 8th edn. Basle: Ciba-Geigy, 1982:73-107.
- 16 Gallen RS. Use of predictive value theory in clinical immunology. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*. 3rd edn. Washington: American Society for Microbiology, 1986:966-71.