

*Quality Assurance Methodology
Refinement Series*

*Comparative Validity
of Three Methods for
Assessment of the
Quality of Primary
Health Care:
Guatemala Field Study*

By:

Jorge Hermida, M.D.

David Nicholas, M.D.

Stewart Blumenfeld, M.D.

Preface

The Quality Assurance Project (QAP) was initiated in 1990 to develop and implement sustainable approaches for improving the quality of health care in less developed countries. QAP has two broad objectives: 1) to provide technical assistance in designing and implementing effective strategies for monitoring quality and correcting systemic deficiencies; and 2) to refine existing methods for ensuring optimal quality health care through an applied research program.

The project's Methodology Refinement component is aimed at developing, refining and validating cost-effective measures for improving the quality of health care. This third report in the Quality Assurance Methodology Refinement Series describes the results of a field study in Guatemala which compared the validity of three quality assessment methods for measuring the adequacy of health worker performance: medical record review, exit interviews with mothers, and direct observation. Readers interested in the issue of assessment method validity and reliability are encouraged to see also a separate QAP Methodology Refinement report entitled, *Comparison of Methods for Assessing Quality of Health Worker Performance Related to Management of Ill Children: Malawi Field Study*. The two studies have important implications for the use of these methods in quality assessments and routine supervision and quality assurance.

The authors wish to express their thanks to Project HOPE and its personnel in Quetzaltenango, Guatemala, and especially the Child Survival Project Director, Dr. Francisco Pineda, for their cooperation with this study and logistic support. The authors want to express their thanks to Lani Marquez for excellent work in editing the original manuscript.

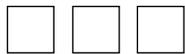
The Quality Assurance Project is funded by the U.S. Agency for International Development under Cooperative Agreement Number DPE-5992-A-00-0050-00 with the Center for Human Services. Collaborating with the Center for Human Services on this project are the Johns Hopkins University School of Hygiene and Public Health and the Academy for Educational Development.

Contents

I. Introduction	7
II. Methods.....	10
III. Results	13
IV. Discussion	16
V. Conclusions	20
Endnotes	21

Comparative Validity of Three Methods for Assessment of the Quality of Primary Health Care: Guatemala Field Study

I. Introduction



The World Health Organization has defined quality health care as services which comply with appropriate national or local standards and are delivered at the required level of care, when needed.¹ Compliance with standards is the key element in this definition, since in order to identify acceptable quality it must be possible to define operationally what specific steps must be taken which together constitute appropriate care. Once such steps are defined, the assessment of quality becomes a task of measuring whether or not the prescribed tasks have been performed.² Discrepancies between actual performance and standard or ideal performance are then identified as quality deficiencies.

Various methods have been used in developed and developing countries' health delivery systems to try to measure whether health services meet acceptable levels of quality. These include record review or audit, interviews with health care providers, written and oral examinations, interviews and focus groups with patients, direct observation of the delivery of services, surrogate patients, retrospective review of adverse outcomes, and simulations³, among others. In developed countries, the most frequent procedure used to assess service quality is the review of medical and other written records.⁴ In less developed countries, however, the use of medical records is often of little use because in many cases the records are quite incomplete, inconsistent or even non-existent, particularly at public ambulatory care facilities. For these reasons, quality improvement teams must usually resort to other means of obtaining data for the evaluation of quality, such as direct observation of care and exit interviews with patients after they have received care.⁵

Although these methods are employed frequently, there has been little empirical research on their validity for measuring the quality of a health worker's performance delivering primary health care services. Validity here is defined as the degree to which a method is able to accurately depict the technical quality of services; in turn, technical quality is operationally defined as the fulfillment by the health workers of key procedures in a standardized way. Yet as interest grows in institutionalizing quality assurance programs, there is an increasing need for accurate information on the validity of different quality assessment methods under diverse conditions. Program managers need to know what is the best method for obtaining data that accurately reflect what has actually happened in the health care delivery process.⁶

The validity of a method used for the assessment of quality is an important issue, since the consequences of using a method of poor validity can seriously undermine any conclusions about quality levels. If a quality assessment method is to be useful, it must tend to accurately reflect the true quality of health workers' performance. Two kinds of errors may occur which threaten the accuracy of assessment methods: first, the method might fail to identify quality deficiencies when they occur, and second, it might incorrectly label a health worker's performance as deficient when indeed the health worker performed adequately. In the former case, the assessment method would have failed to signal a quality problem when one existed, and in the latter, it would have indicated a quality problem where one did not exist. Both types of errors are important and result in resources being misdirected or wasted.

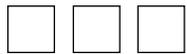
Another important aspect of evaluating the validity of methods relates to costs. Since different methods imply different costs and resources for quality assessment are limited, it is important for program managers to choose a method (or combination of methods) for quality assessment with the lowest cost, at an acceptable level of validity.

This Quality Assurance Project methodology refinement study sought to address the question of assessment method validity by evaluating three methods for quality assessment in a developing country setting: direct observation using checklists, structured interviews with mothers after they received care, and review of medical records completed by the attending health care provider. The validity of each method--i.e., its ability to detect the performance of selected tasks in the care process--was determined by comparing the results obtained through each method with those from an independent

measure of health worker performance that was considered, for the purposes of the study, as reflective of the health worker's "true" performance. The latter measure of performance, known as a "gold standard", was the direct observation of the same service encounter by a specially trained observer with considerable prior experience in the use of observation tools.

The study was carried out in three ambulatory health centers operated by the Ministry of Health in the Department of Totonicapán, Guatemala in May 1994. The centers serve an indigenous rural population in a highlands region. The study was conducted in collaboration with Project Hope, a U.S.-based private voluntary organization which operates child survival projects serving several areas of Guatemala.

II. Methods



Three areas of clinical service were examined to measure service quality: management of acute respiratory infections in children, management of acute diarrhea in children, and family planning counseling to women of reproductive age. For three consecutive weeks, study personnel were stationed in the health centers to collect data on the quality of the clinical care given to all patients requiring the three selected services. In all, encounters were assessed with 74 children with acute respiratory infection, 58 children with acute diarrhea, and 67 mothers who should have received counselling on family planning. The patient encounters involved care delivered by the physician, nurse, and nursing auxiliary which staff each health center; no attempt was made to prescribe which type of health worker should perform each task, although in most cases, it was the physician.

To operationally define what appropriate care would consist of for each of the three services, the study team reviewed the technical norms of the Ministry of Health of Guatemala as well as reference publications of the Pan American Health Organization/World Health Organization.^{7,8} The specific tasks which were selected for each type of service are shown in Figure 1:

Figure 1
Key Health Worker Tasks Assessed by the Three Methods

Acute respiratory infection: <ol style="list-style-type: none">1. The health worker counts the respiratory rate of the child with acute respiratory infection.2. The health worker prescribes an antibiotic if one is indicated.3. The health worker advises the mother on danger signs suggesting pneumonia.
Acute diarrhea: <ol style="list-style-type: none">1. The health worker asks the mother if there was blood in the child's stools.2. The health worker evaluates the dryness of the mucosa in the child's mouth.3. The health worker prescribes oral rehydration salts.4. The health worker advises the mother to continue breastfeeding the child.
Family planning: <ol style="list-style-type: none">1. The health worker advises the mother on the advantages of family planning.

Data on the evaluated clinical encounters were collected by three Guatemalan physicians employed through Project Hope who had similar skill levels as physicians employed by the Ministry of Health. The gold standard observations were performed by two specially trained physicians who had two years

of prior experience applying quality assessment instruments with QAP in a separate initiative of rapid quality assessments of cholera management in Guatemala.

In preparation for their participation in the study, the physician evaluators and “gold standard” evaluators received three days of training directed by an international consultant with extensive field experience in quality assurance. The training included a review of Ministry of Health norms for acute respiratory infection and diarrhea case management, discussion of the data collection instruments and detailed instruction sheets with explicit criteria for each method, simulation exercises, and 8 hours of supervised practice with the instruments in health centers. The gold standard evaluators received an additional day of supervised practice with the international consultant to standardize criteria.

The health personnel in the three study centers were given a very general orientation as to the study’s focus on patient satisfaction and service quality and were not shown the data collection instruments. The providers were unaware of which particular consultations were being observed, since the evaluators accompanied them throughout all consultations. The evaluators applied the observation instruments first, followed by the interview with the mother, and ended with a review of the patient’s clinical record. The field work of all data collectors was supervised daily. Data were processed with the Epi Info software program.

To measure the capacity of the three assessment methods to detect the performance of key tasks by the health worker, the study calculated traditional epidemiological indicators of method error: sensitivity, specificity and predictive value.^{9,10} *Sensitivity* in this case may be defined as the assessment method’s rate of correctly identifying when a task was not performed (a performance failure) when indeed the task was not carried out. *Specificity* is the method’s rate of correctly identifying when a task has been performed (a performance success) when indeed it was performed by the health worker. The *predictive value* of the method is the probability of a performance failure having actually occurred when the assessment method detects a performance failure. In all cases, the “true” occurrence of performance failures and successes was defined as that detected by the gold standard observers.

These measures together characterize how accurate or valid the assessment method is. A method with low sensitivity will often not detect inadequate performance. A method with low specificity will tend to identify adequate performance by a health worker as being inadequate. A method with low predictive value will be unreliable, since only a low proportion of the performance failures it detects will turn out to be real performance failures.

Table 1 illustrates the analyses made of each method, comparing its results to those of the gold standard, for each of the health worker tasks examined.

Table 1
Comparison of Assessment Method vs. Gold Standard for the Performance of Each Task by the Health Worker

METHOD X	GOLD STANDARD		TOTALS
	Task not performed (performance failure)	Task performed (performance success)	
Task not performed (performance failure)	A (true failures)	B (false failures)	A + B
Task performed (performance success)	C (false successes)	D (true successes)	C + D
Total	A + C	B + D	A + B + C + D

The sensitivity of each method was calculated for each discrete task as follows: $A/(A+C)$, that is, the number of cases of performance failure (task not performed by the health worker) detected by the method, divided by the total number of performance failures established by the gold standard.

The specificity of the method was calculated as follows: $D/(B+D)$, that is, the number of cases of performance success (task performed by the health worker) detected by the method, divided by the total number of performance successes established by the gold standard.

The predictive value of the method was calculated as follows: $A/(A+B)$, that is, the proportion of true performance failures among all the failures detected by the method under evaluation.

III. Results



Table 2 shows the sensitivity, specificity and predictive value of the three methods in the assessment of clinical quality (defined as the performance of selected tasks by the health worker) in the care of 56 children with acute diarrhea. It also shows the frequency with which health workers failed to perform the task, as established by the gold standard.

The three methods showed reasonably high levels of sensitivity (generally over 70%) for the detection of failures in the performance of most of the tasks of the health worker.

The sensitivity of direct observation ranges between 72% and 87%, except for the task “evaluation of mouth dryness,” in which the sensitivity of this method was 39%. The specificity of this method was also high—above 83%— except for the task “advice to the mother...,” in which it was only 20%.

Mother’s exit interview was found to have a sensitivity of 69% to 87%, with specificities of over 90%, except for the task “advice to the mother,” in which the specificity was 40%.

Table 2
Sensitivity, Specificity and Predictive Value of the Three Methods for the Assessment of Clinical Quality of Care in 56 Children with Acute Diarrhea

TASK	INDICATOR	ASSESSMENT METHOD		
		Observation	Mother’s interview	Review of medical record
Health worker asks about blood in stools (prevalence of not performing: 67.8%)	<i>Sensitivity</i>	84.2	86.8	97.3
	<i>Specificity</i>	83.3	88.8	33.3
	<i>Predictive value</i>	91.4	94.2	75.5
Health worker evaluates mouth dryness (prevalence of not performing: 67.8%)	<i>Sensitivity</i>	39.4	*	86.8
	<i>Specificity</i>	94.4	*	5.5
	<i>Predictive value</i>	93.7	*	66.0
Health worker prescribes ORS (prevalence of not performing: 32.1%)	<i>Sensitivity</i>	72.2	77.7	77.7
	<i>Specificity</i>	92.1	97.3	89.4
	<i>Predictive value</i>	81.2	93.3	77.7
Health worker advises mother to continue breastfeeding (prevalence of not performing: 69.7%)	<i>Sensitivity</i>	86.9	69.5	100.0
	<i>Specificity</i>	20.0	40.0	10.0
	<i>Predictive value</i>	71.4	72.7	71.8

* The mother was not asked whether the health worker had examined her child’s mouth.

The review of medical records yielded a sensitivity of 77% to 100%, but very low specificities except in regard to the task “prescription of oral rehydration salts,” in which it reached 89%.

Predictive values were in general over 70% for all the tasks assessed and for all three methods. Comparatively, the lowest predictive values were obtained by the review of clinical records. High prevalences of “not performing” in most of these key tasks should be noted. The effect of these prevalence levels on predictive values will be discussed later on.

Table 3 presents sensitivity, specificity and predictive value for the three methods in the detection of tasks performed in the care of 74 children with acute respiratory infection.

Direct observation yielded sensitivity levels above 70%, and high specificity (above 90%). The interview of the mother showed a sensitivity ranging between 57% and 92%, and a specificity of more than 75%.

Review of the medical record showed high sensitivity, but very low specificity except in the task “antibiotic prescribed,” for which it reached the 80% mark. The prevalence of “not performing” these tasks is high, especially for advising the mother.

Table 3
Sensitivity, Specificity and Predictive Value of the
Three Methods for the Assessment of Clinical Quality
of Care in 74 Children with Acute Respiratory Infection

TASK	INDICATOR	ASSESSMENT METHOD		
		Observation	Mother's interview	Review of medical record
Health worker counts child's respiratory rate (prevalence of not performing: 63.5%)	<i>Sensitivity</i>	72.3	57.4	97.8
	<i>Specificity</i>	100.0	92.6	3.7
	<i>Predictive value</i>	100.0	93.1	63.8
Health worker prescribes antibiotic*	<i>Sensitivity</i>	77.7	77.7	83.3
	<i>Specificity</i>	92.8	82.1	80.3
	<i>Predictive value</i>	77.7	58.3	57.6
Health worker advises mother on recognizing danger signs of pneumonia** (prevalence of not performing: 90.5%)	<i>Sensitivity</i>	78.5	92.8	95.2
	<i>Specificity</i>	100.0	75.0	25.0
	<i>Predictive value</i>	100.0	97.5	93.0

* An antibiotic should be prescribed with a diagnosis of pneumonia, ear infection, or bacterial infection of the throat (PAHO/WHO).⁵ This study measured the validity of the three methods for detecting the prescription of an antibiotic to any of the children considered, whatever the diagnosis.

** This action was evaluated only in the cases of mothers whose children were diagnosed with acute lower respiratory infection or pneumonia (n = 46 children).

Table 4 presents the sensitivity, specificity and predictive value of the three methods for detecting whether the health worker counseled the woman on the advantages of family planning.

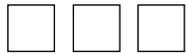
Table 4
Sensitivity, Specificity and Predictive Value of the
Three Methods for the Assessment of Family Planning
Counseling in 67 Mothers

TASK	INDICATOR	ASSESSMENT METHOD		
		Observation	Mother's interview	Review of medical record
Health worker advises mother on advantages of family planning (prevalence of not performing: 79.1%)	<i>Sensitivity</i>	98.1	96.2	100.0
	<i>Specificity</i>	85.7	57.1	14.2
	<i>Predictive value</i>	96.2	89.4	81.5

In general, the three methods evaluated showed very high sensitivity levels for the detection of family planning counseling, but varying levels of specificity: specificity was high for direct observation, intermediate for the interview of the mother, and very low for the review of medical records. Predictive value was high for each of the three methods, with the lowest level for the review of clinical records. Prevalence of “advising the mother” was quite low.

The other aspect of the three assessment methods which was examined in the study was the relative cost of applying each technique. For each application of each method, the evaluators recorded the amount of time spent completing the respective data collection form. The average time required to observe a patient encounter ranged from 7.51 minutes for acute respiratory infection encounters to 8.4 minutes for diarrheal disease encounters; for exit interviews with mothers, from 1.49 minutes to 1.02 minutes; and for record reviews, from 0.83 to 0.75 minutes. The unit cost of supervisor time was costed at US\$0.05 per minute. More significant than the cost of the labor involved in applying each method was the travel cost for a supervisor to visit the health center in order to apply the method. Since this travel cost (estimated at US\$10.89) was the same for each method, the study found little difference in the overall cost of each method.

IV. Discussion



In evaluating the validity of a quality assessment method, a high sensitivity means that the method can detect a large number of real failures of the health worker to perform critical tasks or procedures. Conversely, a low sensitivity would mean a high probability of not detecting many of the health worker's quality failures, erroneously labeling his performance as being of acceptable quality and potentially putting at risk patients' health. On the other hand, a high specificity means that the method is able to recognize when the health worker has performed tasks appropriately. A low specificity indicates that the method takes for performance failures, many of the health worker's actions that are really successes, and risks the waste of resources used unnecessarily in trying to improve quality levels when they are already acceptable.

Predictive value, as a measure of the accuracy of a method of assessment, should be calculated and used when the true prevalence of the condition under scrutiny is known, or if the method is tested on a random sample from the population in which it is to be applied. This is because the predictive value increases (decreases) as the true prevalence of the condition increases (decreases). In the present study, the true prevalence of performance failure in the three health centers was high for most of the tasks, which would tend to elevate the predictive values of the assessment methods. Moreover, the service encounters with women and children which were studied were not selected as a random sample but rather constituted all of the eligible service encounters during the three-week period of study. Thus, the predictive values obtained must be analyzed with caution. It may be argued, however, that the prevalences of the "conditions" under assessment (that is, the performance of specific tasks by health workers) obtained in this study are similar to those obtained in comparable studies done in the past in Latin America.

In general, the sensitivity of the three methods was found to be high for the detection of failures among most of the health workers' tasks. In other words, they were able to detect a high proportion of the tasks not performed by the health worker during the care process. The greatest problem appears to lie in the specificity of the methods, that is, their capacity to recognize quality successes and detect only real failures, instead of labeling as failures performance which is in reality acceptable.

Another way to study variations in the sensitivity and specificity of methods is to examine them in relation to the type of task evaluated. To do this, the analysis grouped the health worker tasks evaluated in four categories: asking

questions, physical examination, prescribing, and advising the patient. Figure 2 shows the sensitivity and specificity of each of the three assessment methods for these task groupings.

Figure 2
Sensitivity and Specificity of the
Three Assessment Methods by Type of Task

TASKS	METHOD ASSESSED		
	Sensitivity	Specitivity	
Asking Questions Health worker asks about blood in feces	84.2	83.3	observation
	86.8	88.8	interview of mother
	97.3	33.3	review of medical record
Physical Examination Health worker evaluates mouth dryness	39.4	94.4	observation
	*	*	interview of mother
	86.8	5.5	review of medical record
Health worker counts child's respiratory rate	72.3	100	observation
	57.4	92.6	interview of mother
	97.8	3.7	review of medical record
Prescribing Health worker prescribes ORS	72.2	92.1	observation
	77.7	97.3	interview of mother
	77.7	89.4	review of medical record
Health worker prescribes antibiotic	77.7	92.8	observation
	77.7	82.1	interview of mother
	83.3	80.3	review of medical record
Advising Patient Health worker advises mother to continue breastfeeding	86.9	20.0	observation
	69.5	40.0	interview of mother
	100	10.0	review of medical record
Health worker advises mother on recognizing danger signs for pneumonia	78.5	100	observation
	92.8	75.0	interview of mother
	95.2	25.0	review of medical record
Health worker advises mother on family planning	98.1	85.7	observation
	96.2	57.1	interview of mother
	100	14.2	review of medical record

* The mother was not asked whether the health worker performed this task.

The review of medical records yielded a high sensitivity, but a very low specificity in the tasks of asking questions, physical examination, and advice to the patient. The only tasks in which this method provided good specificity are those of prescribing, in which the health worker is most often accustomed to writing in the medical record what is being prescribed. The very low specificity of the review of medical records is likely due to their poor quality and is a measure of the infrequency of entries. Since health workers commonly do not record detailed information about service encounters, particularly regarding physical history and patient counseling, record reviews are biased in the direction of including many “false failures” on the part of the health worker.

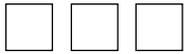
The exit interview with the mother generally yielded very satisfactory results for most of the tasks evaluated. It might be thought that indigenous mothers in a rural area, most of them with very little education, would not be able to identify accurately the tasks performed by a health worker. And yet the sensitivity and specificity of this method were high, except for a low sensitivity to detect the health worker’s failure to count the child’s respiratory rate and a low specificity in the task of advising mothers about continuing breastfeeding and family planning. One explanation for this is that the overall frequency of counseling activities was low. It may also be that the mothers, whose native language is Quiché, did not understand some of the messages given by the health worker in Spanish, and so reported that they had not received them. As such, it is possible that even higher levels of sensitivity and specificity would be obtained in populations with higher levels of schooling.

Observation demonstrated overall the best levels of sensitivity and specificity. Since the gold standard was also a direct observation, the comparison was really one of inter-observer variability. The difference between the regular evaluators and the gold standard observers lay in their training: the gold standard observations were made by professionals with considerable experience in the use of direct observation, while the other evaluators were applying direct observation instruments for the first time outside of the 3-day training. Nonetheless, the tasks being observed were for the most part simple and objective actions, and the observers were given detailed written guidelines establishing the criteria for rating a task as “performed” or “not performed”. The lack of agreement between the gold standard and the regular direct observation of the same task is striking and underscores the need to take inter-observer variability into account when using direct observation as a

quality assessment method. To avoid bias resulting from the lack of agreement between observers, quantitative standardization of observers should be undertaken prior to the actual assessment.

These results suggest that in addition to any variation that may be intrinsic to the method, an important source of variation in the sensitivity and specificity indicators is the skill level of the person applying the method. A skilled observer or interviewer will obtain different results than a less skilled one.

V. Conclusions



The validity of the methods by which data are obtained for evaluation of the quality of health care is an important methodological consideration that must be taken into account when designing quality assessments and routine supervision strategies. No method has universal validity; validity varies with specific aspects of the work setting. Ideally, a method should have high sensitivity and high specificity; one without the other is useless.¹¹ In order to avoid erroneous conclusions, it is important to understand the nature of any biases which each particular quality assessment method has and to select a combination of methods which are most appropriate to the type of tasks to be assessed.

The present study of the quality of primary care in ambulatory care centers serving an indigenous population in the Guatemalan highlands found that the method of reviewing medical records has very low specificity for the detection of the performance of key tasks. Until the quality and completeness of routine medical records improve substantially in developing country settings, the use of record review is not recommended, except for the evaluation of quality in the prescription of drugs.

The study found that the method of the exit interview of the mother has levels of sensitivity similar to those of the two other methods evaluated, and a better specificity than that of medical record review in all tasks evaluated. The advantages of making more use of this method as an alternative to direct observation could include a potentially lower cost and elimination of possible changes in health workers' performance induced by the presence of an observer. Greater use of this method to evaluate quality of care is recommended, as well as further research on the conditions under which exit interview demonstrates high specificity.

While direct observation as a quality assessment method demonstrated the best overall balance of sensitivity and specificity, the study revealed important levels of disagreement between observers. This should be taken into account as a potential source of error when direct observation is used to evaluate quality of care. The potential for inter-observer variability underscores the need for thorough training and quantitative standardization of observers before this method is applied.

Endnotes

- ¹ World Health Organization, Division of Strengthening of Health Services, *Quality assessment and assurance in Primary Health Care*. 1989.
- ² Roemer, M.I., Montoya-Aguilar, C., in World Health Organization, *Quality assurance in primary health care*. Geneva, WHO Offset Publication No. 105, 1988.
- ³ PRICOR, Child Survival Report. *SIMULEX, a methodology to evaluate the quality of service delivery*. Center for Human Services, Bethesda, Maryland, October 1989.
- ⁴ Donabedian, A., "The Quality of Care: how can it be assessed?", in: Graham, N., *Quality Assurance in Hospitals*, Aspen Publishers, Inc., Maryland, 1990.
- ⁵ Nicholas, D.D., Heiby, J.R., Hatzell, T.A., in World Health Organization, *Quality assurance in primary health care*. Geneva, WHO Offset Publication No. 105, 1988.
- ⁶ Saturno, P. Prólogo, in: Nutting, Paul, et. al., *Métodos de evaluación de la calidad en Atención Primaria*. Barcelona, Spain, 1991.
- ⁷ PAHO/WHO, *Atención del niño con infección respiratoria aguda*. PALTEX Series No. 21, Washington, D.C., 1987.
- ⁸ PAHO/WHO, *Manual de tratamiento de la diarrea*. PALTEX Series No. 13, Washington, D.C., 1987.
- ⁹ Lilienfeld, A.M., Lilienfeld, D.E., *Foundations of Epidemiology*, Oxford University Press, New York, 1980.
- ¹⁰ MacMahon, B., Pugh, T., *Principios y métodos de Epidemiología*. La Prensa Médica Mexicana, Mexico City, 1984.
- ¹¹ Bourke, G., Daly, L., McGilvray, J., *Interpretation and uses of medical statistics*. 3rd Edition, Blackwell Scientific Publications, 1985.