Validation of a tool for assessing the quality of pharmaceutical services

Validação de uma ferramenta para avaliar a qualidade de serviços farmacêuticos

Marly Aparecida E. Cosendey [1,2]
Zulmira M. de Araújo Hartz [3]
Jorge A. Z. Bermudez [1]

[1] *Núcleo de Assistência Farmacêutica, Departamento de Ciências Biológicas, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz. Rua Leopoldo Bulhões 1480, Rio de Janeiro, RJ 21041-210, Brasil. bermudez@ensp.fiocruz.br*
[2] *Hospital Universitário Pedro Ernesto, Universidade do Estado do Rio de Janeiro. Boulevard 28 de Setembro 77, Rio de Janeiro, RJ 20551-030, Brasil.*
[3] *Departamento de Epidemiologia e Métodos Quantitativos em Saúde, Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz. Rua Leopoldo Bulhões 1480, Rio de Janeiro, RJ 21041-210, Brasil. zulmira@ensp.fiocruz.br*

Abstract   *This paper presents the validation process for a tool assessing basic pharmaceutical services through an analysis of the implementation of a Basic Pharmaceuticals Distribution Program by the Brazilian Federal government. The process began with the drafting of a theoretical model, based on a state-of-the-art review and allowing the selection of various conceptual dimensions and respective criteria that best represented the construct. The second step involved weighting indicators for the construction of quality scores. Three models were tested for ranking implementation levels, and seven simulations were conducted, determining the score most closely reflecting the selected indicators in two different matrices. The objective was to select the most coherent and consistent version between implementation levels and expected outcomes, while simultaneously enhancing validity of chosen criteria. Testing of the various models and the results obtained showed that augmenting the validity of the study was possible without altering data. This endeavor is justified in understanding the scope and limitations of these measurements and of the choices involved in issues concerning their weighting and interpretation.*
Key words   *Logical Model; Pharmaceutical Services; Assessment*

Resumo   *Neste trabalho apresentamos o processo de validação de um instrumento para avaliar a assistência farmacêutica, por meio da análise de implantação de um programa de distribuição de medicamentos do Governo Federal. Este processo se iniciou com a elaboração de um modelo teórico, baseado na revisão da literatura, que permitiu selecionar as dimensões conceituais e os respectivos critérios que melhor representassem estes constructos. Na segunda fase procedeu-se à ponderação dos indicadores para construção dos escores de qualidade. Foram testados três modelos para a classificação do grau de implantação e sete para a determinação do escore mais adequado com duas matrizes diferentes de indicadores. O objetivo era eleger a que apresentasse a maior coerência e consistência entre o grau de implantação do programa e os efeitos estimados, aprimorando a validade concorrente dos critérios utilizados. Os diversos modelos testados e os resultados obtidos, mostram que é possível se aumentar a validade do estudo, sem alterar os dados coletados, e o esforço se justificando por uma melhor compreensão do alcance e limites das medidas e das escolhas de suas ponderações e interpretações.*
Palavras-chave   *Modelos Lógicos; Serviços Farmacêuticos; Avaliação*

Introduction

Pharmaceutical services consist in a group of activities related to pharmaceuticals and intended to provide support for health care activities required by a community. This also involves storage, stability, and quality control, as well as safeguarding the therapeutic safety and efficacy of the medicines, together with oversight and assessment of their use, obtaining and dissemination of information on the medicines, in parallel to ongoing education of health care practitioners, patients, and communities, in order to ensure rational drug use (Brasil, 1998).

As part of Brazil's National Drug Policy, in 1997 the Federal government drew up its Basic Pharmacy Program in order to rationalize the distribution of a selected group of drugs for treating the most common diseases affecting the Brazilian population. The Program's scope is linked to basic health care services supplied through the National Health System. Striving to cut costs and simplify operations, a standard module or kit was established, consisting of 32 items, identified by generic names and quantified to cover the needs of some 3,000 people for an average period of three months. This kit was acquired from government manufacturers. Laying the foundations for decentralizing basic pharmaceutical services, this program was launched at a time when discussions were underway to define the different components, guidelines, and priorities that were to result in Brazil's National Drug Policy (Bermudez, 2001; Bermudez et al., 2000; MS, 1997).

This paper intends to briefly discuss some concepts concerning the validity of measures in an evaluation tool of pharmaceutical services aimed at analyzing implementation of the Basic Pharmacy Program.

Measurement validity: concepts and classification

Validation is a process estimating the level of compliance of a model or measurement with its respective reality (Palumbo & Oliverio, 1989).

There are several different strategies for establishing measurement validity: construct validity, content validity, and criteria validity (Champagne et al., 1985; Contandriopoulos et al., 1997; Polit & Hungler, 1995; Streiner & Norman, 1989); patent validity (including content validity and consensus validity); as well as criteria validity (Abramson, 1990). Despite these many different classifications of the types of validity, here we assign high priority to that proposed by Champagne et al. (1985); Polit & Hungler (1995); and Contandriopoulos et al. (1997), as these are the most widely used in social welfare and health care programs.

Construct (or *construction*) *validity* covers the relationship between theoretical concepts and their operationalization (Contandriopoulos et al., 1997), meaning the extent to which a tool measures the construction being analyzed (Polit & Hungler, 1995). The main form of construction validation used in this paper was nomological, examining the correlation between various concepts through which the progress of the interrelationships is postulated (Champagne et al., 1985).

Content validity consists of the capacity of the instrument to measure all dimensions of the concepts to be measured (construct). It is a judgment concerning the proportion to which the items selected to measure a theoretical construction properly represent all the dimensions of the concept to be measured (Champagne et al., 1985; Contandriopoulos et al., 1997; Polit & Hungler, 1995). The content validity of a variable can be fine-tuned by breaking down its concept into as many dimensions as possible. To do so, specialist consensus techniques can be used (Champagne et al., 1985). There are two main techniques for determining consensus: the Nominal Group Technique (NGT) and the Delphi Technique. In this study we opted for the NGT. The NGT is a sensitive method used when it is possible to convene the specialists involved. Through this technique, discussion is only permitted at a specific intermediate phase of the process, after individual voting. This technique is used for situations that require group decisions (Abramson, 1990).

Criteria validity refers to the extent to which the scores obtained by an instrument are correlated to certain outside criteria (like a gold standard in diagnosis classification) and is rated as appropriate for use as a valid criterion. The validity of the criterion is ranked as *predictive* when a tool is able to foresee (predict or estimate) some criterion noted at a future moment; it is *concurrent*, *concomitant*, or *simultaneous* when the instrument scores are correlated to some outside criterion measured simultaneously (Abramson, 1990; Champagne et al., 1985; Contandriopoulos et al., 1997; Polit & Hungler, 1995; Streiner & Norman, 1989).

Other approaches based on statistical models can be proposed in order to assess measurement quality. One of them is geometric representation. One advantage of this type of analysis is that the graph can demonstrate not only the relations among the variables that are

assumed to represent the various dimensions, but also shows their potential heterogeneous distribution. The presence of clusters may suggest that the study sample contains two different groups of observation units, for instance. However, one disadvantage of this method is that it limits the number of dimensions to be analyzed and is unable to plot the relative frequency of the study phenomenon in sufficient dimensions. Another constraint is that the dimensions must use the same scale. In view of these limitations, the dimensions are better represented through the use of a ratings vector (observation vector) or scores. The ratings may be added together or can be used as averages or percentages. The vector may also be used to represent the development of the total scores over time. The dimensions of the observation vector variables can be measured on different scales, capable of handling a large number of variables. However, the vector may well be better represented by means of a smaller number of variables by adding scores within each dimension (Miller & Knapp, 1979). In terms of assessing health care service quality, one technique is perhaps the most traditional related to weighting the criteria, constituting Donabedian's diagnostic and therapeutic dimensions (1981, 1986). Score construction allows association between the specialist consensus techniques pre-validating the measurement construct and the weighting used in the second phase (Hartz et al., 1997; Ojeda, 1992; OPS/OMS, 1987). In view of the many dimensions of social welfare and health programs, Kessner et al. (1992) developed a method that selects the conditions representing the target issues, known as "*tracers*". According to Hartz et al. (1995), this method has the advantage of being able to combine the process assessment elements with the service results.

Score construction

The numerical estimations used to sum up the quality of a measurement may be obtained through simple or combined ratings for specific dimensions. When the indicators are used, they can be measured on a nominal, ordinal, or interval scale. For the nominal and ordinal options, the proportion or percentage concept may be used. Either the proportion of desirable responses can be measured separately for the nominal scale, or the proportion within each order can be measured for the ordinal scale. For the interval scale, the most appropriate descriptive measurement or rating is the simple average or the average response within the di-

mension in question. In these three cases, one response may be obtained for each individual question, resulting in a single rating (score) for the entire dimension (Miller & Knapp, 1979).

The score is generally needed when a final result is desired for a specific characteristic or dimension and can be assessed through a series of items. Streiner & Norman (1989) described artifacts that can be used to obtain significant results for a characteristic when the items under assessment are rated at different levels of importance among themselves in terms of dimensions analyzed, all for comparing scores on different metric scales.

Quality assessment of pharmaceutical services

Many countries are developing and implementing national drug policies. Others work on strategies to enhance the quality of pharmaceutical services. However, both processes sometimes lack a systematic approach for evaluating pharmaceutical policies. Hence the need for elaborating tools that permit the effective monitoring of the implementation of national drug policies, in order to evaluate their performance and revise priorities.

After identifying the need for these instruments and at the same time striving for a means of comparison between countries, World Health Organization (WHO) published the *World Drug Situation* in 1988. It presents a great deal of useful information, organized in the form of indicators, the turning point in an endeavor to evaluate the quality of pharmaceutical services on a national basis.

In 1994 WHO published a manual called *Indicators for Monitoring National Drug Policies*, in which a total of 129 indicators for monitoring a wide range of pharmaceutical services in developing countries are described. The book *Rapid Pharmaceutical Management Assessment: An Indicator-Based Approach* (RPM, 1995) aims specifically at standardizing terminology and evaluation methods for drug programs and policies. It introduces a quick evaluation method, a complementary approach to that of WHO.

At present, these manuals are still the reference for the evaluation of drug systems, having guided some of the few studies carried out in Brazil with the purpose of evaluating pharmaceutical services, such as Santich & Galli (1995), Pacheco et al. (1998), Adames (1997), Rozenfeld et al. (1999), and Cosendey (2000).

Methodology

In 1998 the Center for Pharmaceutical Policies at the National School of Public Health (ENSP) of the Oswaldo Cruz Foundation (FIOCRUZ), located in Rio de Janeiro, assessed the implementation of the Basic Pharmacy Program (MS, 1999) in five Brazilian states.

The selection of the process and structure indicators needed to estimate the implementation levels and results of the Program was based on proposals by the WHO and other international institutions for assessing national drug policies (MSH/WHO, 1997; RPM, 1995; WHO, 1993, 1994). The selected indicators were then weighed by a group of specialists. These belonged to the Center for Pharmaceutical Policies at ENSP/FIOCRUZ, a WHO/PAHO Collaborating Center for Pharmaceutical Policies with a multidisciplinary group of ten participants. The nominal group technique was employed. After a database had been built, the indicators were subjected to simulations with different weights and cut-off points in order to upgrade the concurrent validity of the scores obtained. Consequently, the validation of this instrument was based on the coherence between the structure and process indicators with the outcomes indicators and the consistency of the results noted in different contexts. A lack of coherence and/or consistency among these groups of indicators indicated flaws in the assessment tool.

The Basic Pharmacy Program validation process was carried out in three stages, starting with the content validity and the nomological articulation or construction. At this stage, attempts were made to select the macro-dimensions closest to the results that were supposed to be obtained and best reflecting the concept under study. The validation of the construction and content is represented graphically in the logical model (Figure 1).

At a second stage, the content validation was undertaken in phases. Initially, attempts were made to determine the variables needed to represent the concept to be assessed and their relative weight within the construct cluster through agreement among the experts, supported by a review of the Brazilian and international literature. During the second phase, the number of variables was reduced, keeping only those needed to represent the proposed dimensions in the logical model, 24, in order to measure implementation levels, and nine to measure the Program outcomes (Table 1). There is no doubt that this is the most important phase, playing a leading role in con-

structing and fine-tuning these measurement tools.

The final measurement validation in this study was based on criteria validity, meaning the assessment standard used for the implementation levels and the interpretation of the results of the study. This validation is therefore based on the existence of coherence between the structure and process indicators with the outcome indicators and the consistency of the outcomes noted in different contexts, rather than on the equivalency noted among the scores (observation vectors). Consequently, no attempt was made to determine the same number of variables for each dimension assessed or weighed that could assign higher importance to one over another. The lack of coherence or consistency among these groups of indicators showed that there were flaws in the instrument used for the assessment, which led to reviews and adjustments of the validity and content of the instrument, in terms of its variables and weighting.

The Program implementation levels and outcomes were initially weighted by the same group of experts belonging to the Center for Pharmaceutical Policies at ENSP/FIOCRUZ, by way of nominal group technique. After a database had been established, the weights and cut-off points were suggested through model simulations submitted for review, striving to boost the concurrent validity of the scores obtained. However, through empirical observation it was possible to review and enhance the content and construction validity, adjusting the conceptual and operational representations.

The calculation of the implementation level scores was handled in two stages. Initially, the observed and expected values were determined for each dimension, while at a second stage the implementation level was calculated. As not all the indicators used to calculate the implementation level were based on the same scale (some were qualitative) we decided to assign value scores to each indicator according to the results observed, which could be added up to find the value for each dimension as described in Formula 1.
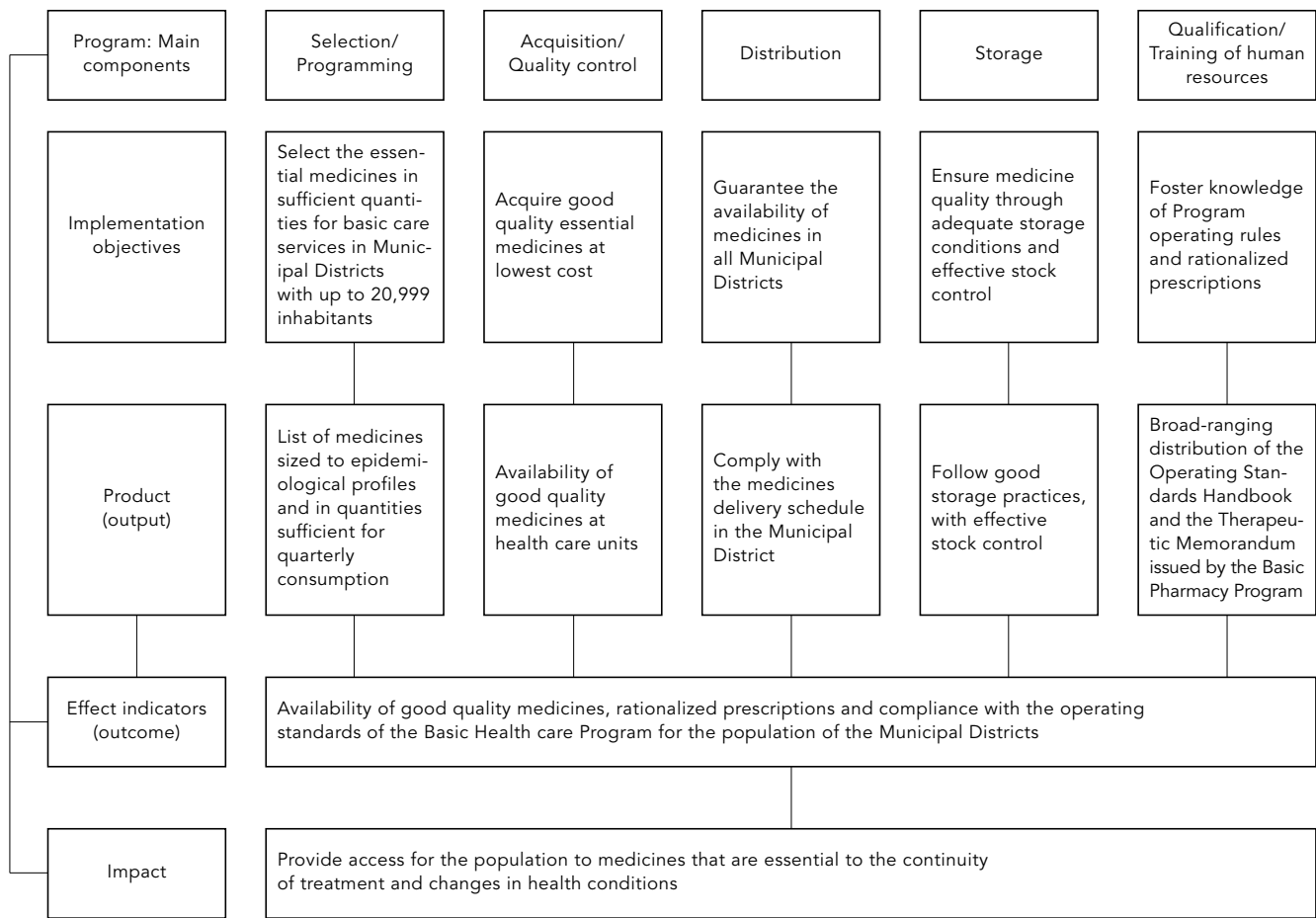
$$Y_i = \Sigma X_{ij}$$

Where:

Y_i = dimension (i), for instance, storage, distribution, etc.

X_{ij} = indicator (j), within dimension (i), for instance, "adaptation of drug storage practices" within the storage dimension.

The score assigned to the expected results always corresponded to the highest possible

Figure 1

Logic model for the Basic Pharmacy Program.

Program: Main components	Selection/ Programming	Acquisition/ Quality control	Distribution	Storage	Qualification/ Training of human resources
Implementation objectives	Select the essential medicines in sufficient quantities for basic care services in Municipal Districts with up to 20,999 inhabitants	Acquire good quality essential medicines at lowest cost	Guarantee the availability of medicines in all Municipal Districts	Ensure medicine quality through adequate storage conditions and effective stock control	Foster knowledge of Program operating rules and rationalized prescriptions
Product (output)	List of medicines sized to epidemiological profiles and in quantities sufficient for quarterly consumption	Availability of good quality medicines at health care units	Comply with the medicines delivery schedule in the Municipal District	Follow good storage practices, with effective stock control	Broad-ranging distribution of the Operating Standards Handbook and the Therapeutic Memorandum issued by the Basic Pharmacy Program
Effect indicators (outcome)	Availability of good quality medicines, rationalized prescriptions and compliance with the operating standards of the Basic Health care Program for the population of the Municipal Districts				
Impact	Provide access for the population to medicines that are essential to the continuity of treatment and changes in health conditions				

value obtained by the observed value and the expected value for each dimension. The implementation level as such was then calculated, as described in Formula 2.

$$\sum_{i=1}^{n} \left(\frac{Yi(O)}{Yi(E)} \right) x100$$

Where:
Yi(O) = value observed for the dimension (i)
Yi(E) = value expected for dimension (i)

The actual measurement of the implementation level was handled at three levels. Initially, differentiated cut-off points were assigned for rating the implementation level; next, different internal scores were assigned to each indicator; finally, the concurrent measurement validation was undertaken. At that point, stage simulations were undertaken for each of the scores, analyzing different cut-off points. Furthermore, simulations were carried out, highlighting some indicators as sentinel program events.

Additionally, three models were tested to determine the most appropriate cut-off point to classify the level of implementation, with seven more used to determine the most appropriate score for the selected indicators, in two different matrixes of indicators. In the first matrix, the drug quality variables that did not adjust to quality ratings for safety were highlighted as sentinel events. They required appropriate investigation. A total of 42 simulations were them performed to elect the option with the highest coherence and consistency between the Program implementation level and the outcomes noted in the various analysis units (Brazilian States)

Table 1

Indicators employed in the implementation analysis. Basic Pharmacy Program.

Program component	Indicator
	B – Structure
Storage	
B12	Adequacy of storage facility in Central County Drug Depot destined for accommodating Program drugs
B13	Adequacy of storage facility in Primary Health Care Facilities destined for accommodating Program drugs
Training of Personnel	
B14	Percentage of Central County Drug Depots that received the Program Operational Procedures Manual
B15	Percentage of Primary Health Care Facilities that publicize the Program Drug List to physicians
	C – Process
Distribution	
C4	Adequacy of drug transportation
C5	Mean time between the first and the second delivery of the Program (in months)
C6	Percentage of complete kits delivered
C7	Mean percentage of out-of-stock period for drugs at the local level
C8	Percentage of drugs available in stock at the Central County Drug Depot
C9	Percentage of drugs available in stock at the Primary Health Care Facility
C10	Availability of drugs in distant counties
C11	Percentage of Central County Drug Depots that receive expired drugs from the Program
Storage	
C12	Percentage of expired drugs in the Central County Drug Depot
C13	Percentage of expired drugs in the Primary Health Care Facility
C14	Mean weighted percentage of inventory variation for drugs stocked at the local level
C15	Mean percentage of inventory registers that correspond to stock check for Program drugs stored at the local level
C16	Percentage of Central County Drug Depots that carry out stock management
C17	Percentage of Primary Health Care Facilities that carry out stock management
C18	Adequacy of drug storage practices at the Central County Drug Depot
C19	Adequacy of drug storage practices at the Primary Health Care Facility
C20	Percentage of Central County Drug Depots that carried out physical inspection of drugs on the last 3 deliveries
C21	Percentage of Primary Health Care Facilities that carried out physical inspection of drugs on the last 3 deliveries
Training of Personnel	
C22	Percentage of Primary Health Care Facilities that dispense drugs through prescription presentation
C23	Percentage of counties that fill out the Drug Supply/Resupply Form
	D – Outcome
Utilization	
D1	Availability of practical information on Program drugs to physicians
D2	Mean number of drugs per prescription
D3	Percentage of prescriptions containing injectable drugs
D4	Percentage of prescriptions containing antimicrobial drugs
D5	Percentage of Program drugs prescribed
D6	Percentage of drugs prescribed by generic name
D7	Percentage of Program drugs dispensed
Political dividends	
D8	Percentage of health-care workers that are informed about the Program
D9	Percentage of patients/users that are informed about the Program

Results and discussion

Models tested to determine the most appropriate cut-off point

Based on the cut-off point to be adopted for ranking the implementation level and the observed effects of the Program, the three models described in Table 2 were tested.

As mentioned previously, the seven models proposed to determine the most appropriate score were tested through the three models described for the implementation level (Table 2) and in two different indicator matrixes, totaling the 42 simulations that were run to discover the score and cut-off point combination with the highest concurrent validity and consistency between the Program implementation level and the outcomes. As a result, when the level of implementation was considered critical, the same rating was expected of the outcome.

Consequently, a combination was obtained between the scores and the cut-off points that proved valid for assessing the implementation level of the Basic Pharmacy Program, which could be useful for assessing other drug programs.

Models tested to determine the most appropriate score

• **Model 1**

Although most of the indicators are expressed in percentages, some are based on different metric scales, making it hard to obtain a single final result (implementation level) or to pin-point the comparison between the observed implementation level and the Program's actual results. Consequently, the decision was made to select a standard for each indicator showing the maximum or minimum level for which the result would be considered ideal. These standards were selected from the literature and prevailing legislation and/or through a consensus of specialists from the Center for Pharmaceutical Policies/FIOCRUZ. A minimum value was assigned to each standard, under which the results obtained would be considered critical. Each indicator consequently had three scores, namely: *value 2*, when the result for the indicator was greater than or equal to that proposed for the standards; *value 1*, when the result of the indicator was between the value proposed for the standard and the minimum acceptable value; and *value 0,* when the result was less than or equal to the proposed minimum, as shown in Table 3.

Table 2

Models analyzed by cut-off points in order to rank the implementation level of the Basic Pharmacy Program.

Rating models	Models		
	A	B	C
Adequate	–	> 75%	–
Acceptable	> 75%	50–75%	≥ 80%
Unsatisfactory	25–75%	25–49%	40–79%
Critical	< 25%	< 25%	0–39%

Table 3

Example with the values assigned to the indicators in model 1.

Quantitative indicators	Results observed %	Value range %	Value obtained	Maximum value
Storage				
Adaptation of drug storage practices at the local level and distribution centers.	49.9	≥ 90 = 2 50–89 = 1 < 50%= 0	0	2
Percentage of drugs beyond their expiration date at the local level and in distribution centers.	0.6	0 = 2 > 0 = 0	0	2
Percentage of drugs physically inspected during the last three deliveries at the local level and at the distribution centers.	0	≥ 90 = 2 50–89 = 1 <50 = 0	0	2
Subtotal	–	–	0	6

The values established for the standards may rate as the maximum or minimum acceptable. When the proposed standard for the indicator is rated zero for the acceptable value, the indicator was treated as a dichotomous variable, where zero value was assigned 2 points and any value higher than zero was scored as zero.

This weighting scheme did not prove adequate due to the fact that some States presented a zero implementation level (Table 4), suggesting that no aspects of the Program had been implemented, not reflecting the real situation.

We then decided to review the indicators used to compose the dimensions, altering the matrix initially proposed. We consequently obtained the results in Table 5, which were tested in relation to the cut-off points.

These results proved neither coherent nor consistent, with the second model then prepared, as discussed below.

• **Model 2**

In this model, the standards stipulated for each indicator were replaced by value ranges determined by quartiles, other than the indicators that behaved as dichotomous variables that were treated in the same way as in the previous model, although assigned different scores. Consequently, the value assigned to each indicator was calculated, as shown in Table 6.

In this model, the score assignment scale was reversed for some indicators. For instance, for the "average weighted inventory variation percentage" indicator, the lower the variation, the higher the score assigned. An example of this model is given in Table 7.

The result obtained for the implementation level was then tested with regard to the cut-off points, obtaining the results given in Table 8.

These results proved more adequate than those for the previous model, particularly for rating the implementation level using the A and C Model. Nevertheless, we decided to test a third model.

• **Model 3**

This model did not rank the results obtained by value ranges, but rather sought to use a decimal scale obtained for the indicator in order to calculate the implementation level. Consequently, indicators giving results in percentages were assigned values from zero to 10, corresponding to one-tenth of the result obtained. For instance, an indicator with a result of 73% was assigned the value of 7.3.

The indicators that behaved as dichotomous variables were assigned a value of 10 when the observed value fell within the expected range and zero for the opposite case. For instance, for the indicator "percentage of prescriptions containing injectable drugs", if the observed value was less than or equal to 1 (one tenth of 10%), the indicator was assigned a score of 10; if it was greater than 1, it was assigned a score of zero.

As in the previous model, the score assignment scale was reversed for some indicators.

Table 4

Implementation level by State – model 1 – version 1.

State/Case	Implementation level %	Effect %
Acre	0	14
Amazonas	0	29
Goiás	17	50
Pernambuco	8	64
Rio de Janeiro	25	36

Table 5

Implementation level – model 1 – version 2.

Case	Model 1A				Model 1B				Model 1C			
	Implementation level %		Effect %		Implementation level %		Effect %		Implementation level %		Effect %	
Acre	12	C	14	C	12	C	14	C	12	C	14	C
Amazonas	12	C	29	I	12	C	29	I	12	C	29	C
Goiás	25	I	50	I	25	I	50	A	25	C	50	I
Pernambuco	19	C	64	I	19	C	64	A	19	C	64	I
Rio de Janeiro	33	I	36	I	33	I	36	I	33	C	36	C

C = Critical; I = Unsatisfactory; A = Acceptable.

For instance, for the indicator "weighted inventory variation average percentage", the scale varies from (-)10 to 0, meaning that the lower the variation, the higher the score.

An example of this model is given in Table 9.

The implementation level results were tested for the cut-off points described above, with the results shown in Table 10.

This model proved coherent only when the implementation level was calculated in accordance with Model A.

Table 6

Determination of the value assigned to each indicator.

% of the maximum value obtained	Dichotomous variables	Score assigned
< 25	Result outside stipulated range	1
25–49	–	2
50–74	–	3
≥ 75	Result within stipulated range	4

Table 7

Example with the values assigned to the indicators in model 2.

Quantitative indicators	Results observed %	Value range %	Value obtained	Maximum value
Storage				
Adaptation of drug storage practices at the local level and distribution centers	49.9	> 75 = 4 > 50–75 = 3 > 25–50 = 2 ≤ 25 = 1	2	4
Percentage of drugs beyond their expiration date at the local level and in distribution centers.	0.6	0 = 4 > 0 = 1	1	4
Percentage of drugs physically inspected during the last three deliveries at the local level and at the distribution centers.	0	> 75 = 4 > 50–75 = 3 > 25–50 = 2 ≤ 25 = 1	1	4
Average weighted percentage inventory variation for medicines stored at the local level and at the distribution centers	0*	> 75 = 1 > 50–75 = 2 > 25–50 = 3 ≤ 25 = 4	1	4
Subtotal	–		5	16

* No stock control was undertaken, which was why it was assigned the lowest possible value.

Table 8

Implementation level – model 2.

Case	Model 2A				Model 2B				Model 2C			
	Implementation level %		Effect %		Implementation level %		Effect %		Implementation level %		Effect %	
Acre	41	I	50	I	41	I	50	A	41	I	50	I
Amazonas	47	I	54	I	47	I	54	A	47	I	54	I
Goiás	50	I	71	I	50	A	71	A	50	I	71	I
Pernambuco	53	I	75	I	53	A	75	A	53	I	75	I
Rio de Janeiro	69	I	57	I	69	A	57	A	69	I	57	I

C= Critical; I = Unsatisfactory; A = Acceptable.

Table 9

Example with the values attributed to the indicators in model 3.

Quantitative Indicators	Results observed %	Value obtained	Maximum value
Storage			
Adaptation of drug storage practices at the local level and distribution centers	49.9	4.99	10
Percentage of drugs beyond their expiration date at the local level and in distribution centers.	0.6	-0.06	0
Percentage of drugs physically inspected during the last three deliveries at the local level and at the distribution centers.	0	0	10
Subtotal	–	4.93	40
Effect			
Average number of medicines per prescription	2.4	0	10
Percentage of prescriptions containing injectable drugs	29.0	0	10
Subtotal	–	0	20

Table 10

Implementation level – model 3.

Case	Model 3A				Model 3B				Model 3C			
	Implementation level %		Effect %		Implementation level %		Effect %		Implementation level %		Effect %	
Acre	25	I	49	I	25	I	49	I	25	C	49	I
Amazonas	32	I	59	I	32	I	59	A	32	C	59	I
Goiás	36	I	75	I	36	I	75	A	36	C	75	I
Pernambuco	32	I	72	I	32	I	72	A	32	C	72	I
Rio de Janeiro	51	I	65	I	51	A	65	A	51	I	65	I

C= Critical; I = Unsatisfactory; A = Acceptable.

• **Models 4, 5, 6, and 7**

Models 4, 5, and 6 were variations on Model 3 in terms of the dichotomous variables: in Models 4 and 5 the only difference was that the scores assigned to the dichotomous variables were 1 and 2 and 1 and 4, respectively, rather than 0 and 10 as in Model 3.

In Model 6, a value for the indicator greater than the maximum value stipulated as the standard for the dichotomous variable was assigned a negative value corresponding to the number of surplus points. If the result fell within the value range stipulated for the indicator,

this value was computed as a positive point. For instance, for the indicator "average number of drugs per prescription", the stipulated standard was 2, so if the observed value was 1.4, this value would be considered less than 2. However, if the observed value was 2.4, the score for this indicator would be (-)0.4.

Model 7 can be better understood through an example. The "average number of drugs per prescription" indicator is assigned a standard of 2 with the study obtaining a value of 2.3. Consequently, a score of 10 was assigned to the standard and the score for the value was 2.3, obtained through the reverse rule of three, in

this example reaching 8.29. Similarly, all indicators behaving as dichotomous variables were processed in the same way.

These Models were tested similarly to their predecessors, showing no coherence or consistency between the implementation levels found and the effects observed.

Conclusions

The construction of a summary rating ultimately requires value judgments (PNUD/IPEA/FJP/IBGE, 1998). Based on tests carried out with the various Models presented above, we decided to work with Model 2 to determine the scores. Models A and C displayed adequate combinations for ranking the implementation level. However, we opted for Model C, which had already been used to assess the Mother and Child Program, as the classification recommended by PAHO/WHO (1987), also used by Ojeda (1992) and Hartz et al. (1997), due to the fact that Model A has a very broad range for the *unsatisfactory* classification (25-75%).

The entire procedure for enhancing the validity of the study did not in any way alter the observations or the data collected. All efforts were focused on obtaining a better understanding of the limits of the quantitative measurements and the choices (always flawed) of their weighting and interpretation.

The various models tested after the construction of the Logical Model, the validation, and the assessment of the results all demonstrate the complexity of assessing pharmaceutical services programs, highlighting the need to seek consistency between implementation levels and observed outcomes, based on the basic analytical units selected.

References

ABRAMSON, J. H., 1990. *Métodos de Estudio en Medicina Comunitaria: Una Introducción a los Estudios Epidemiológicos y de Evaluación*. Madrid: Ediciones Díaz de Santos.

ADAMES, L. A. B., 1997. Medicamentos essenciais: Gestão nos serviços públicos de saúde fundamentada no quadro nosológico. *Divulgação Saúde em Debate*, 18:34-42.

BERMUDEZ, J., 2001. Au Brésil, le triomphe des génériques. *Biofutur,* 210:40-42.

BERMUDEZ, J. A. Z.; EPSZTEJN, R.; OLIVEIRA, M. A. & HASENCLEVER, L., 2000. *O Acordo Trips da OMC e a Proteção Patentária no Brasil: Mudanças Recentes e Implicações para a Produção Local e o Acesso da População aos Medicamentos*. Rio de Janeiro: Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz.

BRASIL, 1998. Portaria nº 3.916/GM de 30 de outubro de 1998, dispõe sobre a aprovação da Política Nacional de Medicamentos. Brasília, DF: *Diário Oficial da República Federativa do Brasil*, nº 215E, seção I, p. 18, 10 nov.

CHAMPAGNE, F.; CONTANDRIOPOULOS, A. P. & PINEAULT, R., 1985. Un cadre conceptuel pour l'évaluation des programmes de santé. *Révue d'Épidemiologie et Santé Publique,* 33:173-181.

CONTANDRIOPOULOS, A. P.; CHAMPAGNE, F.; POTVIN, L.; DENIS, J. L. & BOYLE, P., 1997. *Saber Preparar uma Pesquisa: Definição, Estrutura e Financiamento*. 2ª Ed. São Paulo: Editora Hucitec.

COSENDEY, M. A. E., 2000. *Análise da Implantação do Programa Farmácia Básica: Um Estudo Multicêntrico em Cinco Estados do Brasil*. Tese de Doutorado, Rio de Janeiro: Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz.

DONABEDIAN, A., 1981. Criteria, norms and standards of quality: What do they mean? *American Journal of Public Health,* 71:409–412.

DONABEDIAN, A., 1986. Criteria and standards for quality assessment and monitoring. *Quality Review Bulletin,* 12:99-108.

HARTZ, Z. M. A.; CHAMPAGNE, F.; CONTANDRIOPOULOS, A. P. & LEAL, M. C., 1997. Avaliação do programa materno-infantil: Análise de implan-

tação em sistemas locais de saúde no nordeste do Brasil. In: *Avaliação em Saúde: Dos Modelos Conceituais à Prática na Análise da Implantação de Programas* (Z. M. A. Hartz, org.), pp. 29-47, Rio de Janeiro: Editora Fiocruz.

HARTZ, Z. M. A.; CONTANDRIOPOULOS, A. P.; CHAMPAGNE, F. & LEAL, M. C., 1995. Uso de traçadores para avaliação de qualidade da assistência à criança: O controle da doença diarreica e das infecções respiratórias agudas em dois centros de saúde. *Revista do IMIP*, 9:35-50.

KESSNER, D. M.; CAROLYN, E. K. & SINGER, J., 1992. Evaluación de la calidad de la salud por el método de los procesos trazadores. In: *Investigaciones sobre Servicios de Salud* (K. L. White, org.), pp. 189-193, Washington, DC: Organización Panamericana de la Salud.

MILLER, M. C. & KNAPP, R. G., 1979. *Evaluating Quality of Care: Analytic Procedures – Monitoring Techniques.* London: Apsensystens Corporation.

MS (Ministério da Saúde), 1997. *Farmácia Básica: Programa 1997/98.* Brasília: MS.

MS (Ministério da Saúde), 1999. *Relatório de Avaliação do Programa Farmácia Básica.* Brasília: MS. (mimeo.)

MSH (Management Sciences For Health)/WHO (World Health Organization), 1997. *Managing Drug Supply: The Selection, Procurement, Distribution, and Use of Pharmaceuticals.* Bloomfield: Kumarian Press.

OJEDA, N. S., 1992. Evaluación de servicios de salud maternoinfantil en América Latina. *Foro Mundial de la Salud*, 13:140-143.

OPAS (Organização Pan-Americana da Saúde)/OMS (Organização Mundial da Saúde), 1987. *Condições de Eficiência dos Serviços de Atenção Materno Infantil.* Washington, DC: OPAS.

PACHECO, F.; AGUIAR, M. G. G. & QUEIROZ, A. M., 1998. *Diagnóstico da Assistência Farmacêutica no Estado da Bahia.* Salvador: Instituto de Saúde Coletiva, Universidade Federal da Bahia. (mimeo.)

PALUMBO, D. J. & OLIVERIO, A., 1989. Implementation theory and the theory-driven approach to validity. *Evaluation and Program Planning*, 12: 337-344.

PNUD (Programa das Nações Unidas para o Desenvolvimento)/IPEA (Instituto de Pesquisa Econômica Aplicada)/FJP (Fundação João Pinheiro)/IBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 1998. *Desenvolvimento Humano e Condições de Vida: Indicadores Brasileiros.* Brasília: PNUD/IPEA/FJP/IBGE.

POLIT, D. F. & HUNGLER, B. P., 1995. *Fundamentos de Pesquisa em Enfermagem.* 3ª Ed. Porto Alegre: Artes Médicas.

ROZENFELD, S.; PORTELA, M. C.; BERMUDEZ, J. A. Z. & MOSEGUI, G. B. G., 1999. Distribuição de medicamentos: Proposta para avaliação de impacto. In: 1º Congresso de Ciências Farmacêuticas do Rio de Janeiro, *Anais*, pp. 93-94. Rio de Janeiro: Conselho Regional de Farmácia do Estado do Rio de Janeiro.

RPM (Rational Pharmaceutical Management Project), 1995. *Rapid Pharmaceutical Management Assessment: An Indicator-Based Approach.* Washington, DC: Kumarian Press.

SANTICH, I. R. & GALLI, A., 1995. *La Asistencia Farmacéutica en el Sistema Único de Salud de Ceará: Elementos para su Transformación.* Fortaleza: Departamento de Assistência Farmacêutica, Escola de Saúde Pública.

STREINER, D. L. & NORMAN, G. R., 1989. *Health Measurement Scales.* New York: Oxford University Press.

WHO (World Health Organization), 1988. *The World Drug Situation.* Geneva: WHO.

WHO (World Health Organization), 1993. *How to Investigate Drug Use in Health Facilities: Select Drug Use Indicators.* Geneva: WHO.

WHO (World Health Organization), 1994. *Indicators for Monitoring National Drug Policies: A Practical Manual.* Geneva: WHO.