

Use of the Nominal Group Technique and the Delphi Method to draw up evaluation indicators for strategies to deal with violence against children and adolescents in Brazil

Uso da Técnica Grupo Nominal e do Método Delphi para a elaboração de indicadores de avaliação das estratégias de enfrentamento da violência contra crianças e adolescentes no Brasil

Suely Ferreira Deslandes ¹
 Corina Helena Figueira Mendes ²
 Thiago de Oliveira Pires ³
 Daniel de Souza Campos ⁴

^{1,3,4} Centro Latino-americano de Estudos de Violência e Saúde (CLAVES). Instituto Fernandes Figueira (IFF). Fundação Oswaldo Cruz (FIOCRUZ). Av Rui Barbosa, 716. 2º andar. Flamengo. Rio de Janeiro, RJ, Brasil. CEP: 21.040-361.
 E-mail: deslandes.s@gmail.com

² Instituto de Pesquisas Evandro Chagas (IPEC). Fundação Oswaldo Cruz (FIOCRUZ). Rio de Janeiro, RJ, Brasil.

Abstract

Objectives: the aim of this study is to present a method for the production and selection of indicators to evaluate and/or monitor strategies to: a) prevent violence and promote protective family and community relations; b) provide care for victims of such violence and their family members; c) upgrade the recording and reporting of such violence; d) guarantee the rights of child and adolescent victims; and e) ensure due prosecution of perpetrators.

Methods: consensus-generating participatory methodologies were used (Delphi Method and Nominal Group Technique).

Results: 113 indicators were produced, with 27 and 91 indicators selected in scenarios with different scores.

Conclusions: the consensus methodologies were adequate for the selection and validation of evaluation indicators, but criteria need to be established for selection among the indicators adopted.

Key words *Family relations, Indicators, Violence, Public policy, Child, Adolescent*

Resumo

Objetivos: o estudo apresenta o processo de produção e seleção de indicadores para avaliar e/ou monitorar as estratégias de: a) prevenção e promoção de relações familiares e comunitárias protetoras; b) atenção às vítimas destas violências e aos seus familiares; c) qualificação de registros e notificação destas violências; d) garantia dos direitos das crianças e adolescentes vitimizados; e) responsabilização dos autores destas violências.

Métodos: foram empregadas metodologias participativas geradoras de consenso (Método Delphi e Técnica Grupo Nominal).

Resultados: foram produzidos 113 indicadores e selecionados 27 e 91 indicadores em cenários com escores diferentes.

Conclusões: as metodologias de consenso se mostraram adequadas para a eleição e validação de indicadores para avaliação, porém demandam que se estabeleçam critérios de seleção dentre os indicadores aprovados.

Palavras-chave *Relações familiares, Indicadores, Violência, Políticas públicas, Criança, Adolescente*

Introduction

The present article reports on a methodological experience involving the drawing up of indicators for performance evaluation of Brazilian municipalities in dealing with violence against children and adolescents, using consensus-generating participatory methodologies. This experience was part of a study conducted in partnership with the United Nations Children's Fund (UNICEF), entitled "Indicators for the Evaluation of Strategies to Deal with Domestic Violence and Sexual Exploitation of Children and Adolescents".

The article focuses on the methods used to produce and select indicators for evaluation and/or monitoring of the way these kinds of violence are dealt with, in the following areas: a) prevention of domestic violence and sexual exploitation, and promotion of protective family and community relations; b) care for victims of domestic violence and sexual exploitation and their family members; c) improving the reporting and recording of incidents of domestic violence and sexual exploitation; d) defending the rights of children and adolescents who are victims of domestic violence and sexual exploitation; and e) due prosecution of the perpetrators of domestic violence and sexual exploitation against children and adolescents.

These areas topics cover the full range of protection and care that public policies provide victims in Brazil. These initiatives seek to minimize or eradicate such forms of violence, the real magnitude of which is still unknown in Brazil. Data on these kinds of violence tend to be an underestimate its prevalence. The Information System for Childhood and Adolescence, under the Special Secretariat for Human Rights, has a database fed discontinuously by the country's Tutelary Councils. From January 1999 to February 2010, a total of 1,000,621 violations of rights were recorded, of which 466,941 involved violation of the right to a family and community contact (the main item under which cases of domestic violence are tabulated).¹ The "Dial 100" reporting service recorded 90,407 cases of violence between May 2003 and May 2009, of which 31% were cases of sexual violence and 39.8% cases involving sexual exploitation.² A study conducted by the University of Brasilia and UNICEF showed the proven existence of cases of sexual abuse in 937 of the 5,562 Brazilian municipalities.³

Owing either to the low capacity to record the true magnitude of the problem or to the severity of consequences for the health and citizenship of child and adolescent victims, the time has come to demand

an agenda for the evaluation of the ability of local government initiatives in Brazil to deal with these issues.

Consensus methodologies have been identified as the most appropriate for producing evaluation criteria or indicators in the absence of acknowledged and validated scientific and technical parameters.⁴⁻⁶ In addition to being technically recommended, the choice of participatory methods and techniques not only constitutes a methodological strategy, but also expresses an ethical and heuristic principle in relation to the opinions, expectations, knowledge, and experiences of those who study and are active in dealing with domestic violence and sexual abuse and the exploitation of children and adolescents. With their different peculiarities and competences, such academics and service providers are viewed as partners in the production of knowledge.

The initiative of collectively constructing and selecting a set of indicators to evaluate the municipal strategies is a pioneering task in Brazil. The development of indicators may provide support for summative and formative evaluations, in addition to monitoring and social control of action and allowing diagnoses to be carried out that suggest how oversights and shortcomings might be overcome. This approach is especially appropriate for diagnostic evaluations and so-called thematic evaluations, which focus on a global or regional level and are applied in the following cases: a) theme-related; b) evaluation of a policy area from the perspective of its impacts or development or related to the evaluation of various policies with their combined effects (policy mix); c) evaluation to assist a given sector (health, education, etc.).⁷

Methods

The construction and selection of indicators followed five complementary stages: Stage 1: definition of the framework for the creation of evaluative questions, judgment criteria, and indicators by the research team; Stage 2: development of indicators by experts using the Nominal Group Technique (NGT) and evaluation of indicators developed by the research team; Stage 3: review of the set of indicators by the research team; Stage 4: selection of experts in all the regions of Brazil, submission of indicators to this group (Delphi Method), and analysis of the first expert consultation; and Stage 5: second expert consultation (Delphi Method) and final analysis. As this article reports on the methodological experience, each of these phases will be described in detail in the results section.

NGT is characterized by the presence of subjects in a collective meeting, when the participants report their opinions and proposals in writing and proceed to discuss them with the group. The dynamic unfolds through the work of a facilitator, who conducts the debate among the experts, in the form of a structured meeting, generally involving 9 to 12 participants.⁸

The guest participants are considered "experts" in the broad sense of the term, ranging from academic expertise to individuals whose life experience is significant for the issue at stake.⁹

The experts were selected through a network of key informants, with a telephone contact followed by an official personal invitation. The experts were required to be researchers, managers, or professionals with acknowledged work in one of the five mentioned areas.

The Delphi Method, developed in the 1950s in the United States, originally aimed to gather forecasts on international political and military issues.¹⁰ Its use is still heavily associated with the political sciences and is defined as a structured expert consultation, seeking convergence of analyses of future scenarios.¹¹ However, its use has since spread to other fields of knowledge, in particular health policy and other public policy areas.¹⁰

Despite variations, the Delphi method has some characteristic stages:⁸ a) during the first round, the subjects considered to be experts, are invited to give their opinions on the problem in hand, based on a prior consultation; b) during the second round, a questionnaire with the proposals and questions raised is elaborated and submitted to the group. Individual participants then answer the questionnaire, recording their agreement or disagreement with each proposed item. They can also be asked to score or rank each item's importance within the set. These answers are tabulated and can be resubmitted

to the group. In the current study, the first round was based on NGT and the study thus began with the second round.

It is important to note that such forms of consensus are transient and subject to the contextual and discursive historicity of each area, but can nevertheless provide a basis for validating the evaluation.

The work was conducted over the course of eight months, from July 2007 to March 2008.

The Delphi consultation included experts dealing with rights and representatives of government agencies and legitimate nongovernmental organizations and universities from the five regions of Brazil.

Selection of the group of experts was supported by the regional offices of UNICEF, which submitted lists of individuals from various fields. The 'snowball' technique was used to select this group of experts. As shown in Table 1, a total of 746 experts were consulted, of whom 164 (22%) responded to the first Delphi consultation and 120 responded to the second (73.2% of the group that responded). The experts were predominantly from the Southeast Region (both those consulted and those that responded). There were two reasons for this: 1) the study covered professionals that had participated in the NGT and 2) there is known to be a concentration of experts in this region. The Northeast Region was second, with a significant contingent of experts in the field. In the group that actually responded to the first Delphi consultation, there were representatives from the Federal District (Brasilia) and all the States (except Amazonas, whose experts did not respond to the consultation). The request was made by e-mail.

The study was approved by the Ethics Committee of the Instituto Fernandes Figueira (Number CAAE 0035.0.008.000-07) of the Oswaldo Cruz Foundation.

Table 1

Distribution of experts consulted and those that responded, by Brazilian region, 1st and 2nd Delphi consultation, 2007.

Regions	Consulted	Responded 1 st		Responded 2 nd	
		n	%	n	%
Central-West	76	12	15.8	11	91.7
Northeast	196	40	20.4	27	67.5
North	90	16	17.8	12	75.0
Southeast	286	79	27.6	56	70.9
South	98	17	17.3	14	82.3
Total	746	164	22.0	120	73.2

Results

The methodological experience and its results

The main purpose of Stage 1 was to establish a framework based on consultation of the principal public policies and technical norms in the areas of justice, health, human rights, and social work, in so far as these deal with domestic violence and sexual exploitation in Brazil. This collection of documents was selected according to the following criteria: whether the policies were current and in force; nationwide scope and enforcement; and diversity of authorship and origin according to government sectors and agencies.

State and municipal plans and laws were not included, as these establish local strategies for policy enforcement, thereby precluding generalized appli-

cation.

The content of each document was examined in detail, focusing on the goal and purposes of each policy or proposal. In other words, we examined what was lay closest to the framework of each proposal.¹² A cross-sectional analysis of the documents was performed along thematic lines, once the principal evaluation questions and judgment criteria for the elements that were considered necessary for good performance by the municipality had been drawn up.

The evaluation questions give the evaluation a direction and focus. The criteria define the characteristics of what can be considered successful implementation. The judgment criteria specify the level that the intervention must reach in order to be considered successful (Table 2).

A group of 77 indicators were initially drawn up,

Table 2

Framework of evaluation questions and criteria for the elaboration of indicators proposed by the research team.

Lines of analysis	Evaluation questions	Judgment criteria
1. Prevention of domestic violence and sexual exploitation and promotion of protective family and community relations	Does the municipality invest in strategies for prevention and promotion of protective family and community relations? Are such strategies based on research or diagnoses? Do they show adequate coverage? Are they part of an inter-sector action network?	<ul style="list-style-type: none"> • Investment in the elaboration of diagnoses aimed at supporting prevention. • Capacity to supply actions. • Capacity for coverage by actions. • Investment in the creation and consolidation of networks.
2. Attending to victims of domestic violence and sexual exploitation and their family members	Does the municipality offer adequate, sufficient and networked care for child and adolescent victims and their family members? Does the municipality invest in training professionals to provide quality care?	<ul style="list-style-type: none"> • Installed capacity of the health system for attending child and adolescent victims of violence sexual. • Existence of intra- and inter-sector linkage. • Continuous supply of safeguards/ protective measures in extreme situations. • Supply of care for perpetrators of domestic violence. • Qualification and training of professionals.
3. Improving reporting and recording of domestic violence and sexual exploitation	Has the municipality consolidated a management system for reporting and recording the events?	<ul style="list-style-type: none"> • Supply of training in reporting and recording. • Existence of resources (material and human) in the management of information. • Existence of data banks and systems.
4. Defending the rights of child and adolescent victims of domestic violence and sexual exploitation	Has the municipality taken measures to strengthen the System for Safeguarding the Rights of Children and Adolescents?	<ul style="list-style-type: none"> • Consolidation of policies to strengthen the System for Safeguarding the Rights. • Existence of municipal coping strategy for violence and sexual exploitation. • Support for creation and operation of Tutelary Councils.
5. Due prosecution of perpetrators of domestic violence and sexual exploitation against children and adolescents	Has the municipality taken steps to prosecute/ hold accountable the perpetrators of violence?	<ul style="list-style-type: none"> • Ability to offer specialized legal institutions and public safety.

of which 28 were specific to sexual abuse and exploitation. The indicators were classified as follows:

Area 1: prevention of domestic violence and sexual exploitation and promotion of protective family and community relations – 14 indicators were created; Area 2: attending to victims of domestic violence and sexual exploitation and their family members – 28 indicators; Area 3: improving the reporting and recording of domestic violence and sexual exploitation - 19 indicators; Area 4: defending the rights of child and adolescent victims of domestic violence and sexual exploitation – 13 indicators; Area 5: due prosecution of perpetrators of domestic violence and sexual exploitation against children and adolescents – 3 indicators.

During Stage 2, the Nominal Group Technique (NGT) was applied for critical examination and validation (by experts from the public sector and civil society) of the indicators created by the research team. This procedure was also helped them to draw up new indicators.

A total of 58 participants were invited, and a total of 42 experts came to participate in the total of six meetings, between August and October 2007. One meeting was held for each thematic area, plus an additional session focusing exclusively on the issue of sexual exploitation.

The group of experts included members from the following areas: Health, Social Work, Education, Rights Council, Minors' Court, Office of the State Public Prosecutor, Office of the Public Defender, Tourism, Public Security, Universities, and civil society organizations.

The NGT dynamic involves two Stages, both during the same meeting, which lasted approximately six hours. Put briefly, application of the technique followed the five basic steps outlined in the literature.¹³⁻¹⁵

First step: the question was posed to the group clearly and precisely, explaining the study's objectives, the concept of an indicator, and the proposal for the work session's dynamics.

Permission was requested to record the session and a free and informed consent form was presented for participation in the study.

The participants were then asked to write (individually) at least one proposed indicator related to the theme discussed by the group.

Second step: the participants then read their proposed indicators, which were transcribed and shown on the screen. Each expert then presented the indicators created, without interruption. This is referred to as a "round robin" recording format,

created in such a way that one participant does not influence the other in a chain of answers.

Third step: after the presentation, the individual participants presented arguments in support of their proposals. Interruptions by group members were permitted and limited by the facilitator in order to ensure each participant had enough time to speak. During this phase, some ideas were grouped, and repeated ideas were consolidated by means of negotiation among the experts, thus avoiding duplication of proposals. During this stage, the initial list underwent changes, and the group drew up a second version.

Fourth step: this stage involved voting on the descriptors and indicators, and the group assigned a Likert-type score matrix on the importance of these indicators, establishing an ordinal ranking of the four most relevant indicators (0 - no importance whatsoever; 1 - virtually no importance; 2 - not very important; 3 - medium importance; 4 - important; 5 - very important; 6 - highly important).^{9,16}

The experts produced 147 indicators, classified as follows in the thematic areas: Area 1 - 25; Area 2 - 25; Area 3 - 25; Area 4 - 37; Area 5 - 35.

Fifth step: after this round involving production of the indicators by the experts, the 77 indicators produced by the team were discussed and criticized by the participants and then submitted to the same score matrix (Likert scale) that had been used for the previous evaluation stage.

Stage 3 involved the review and reorganization of the set of indicators by the research team. Indicators were only retained if they achieved a percentage of 70% or greater, grouping items 5 and 6 together, with the maximum score matrix values.

Based on this criterion, 70 of the 77 indicators produced by the research team (90.9%) and 85 of the 147 indicators produced by the experts (57.8%) were approved, giving a total of 155 indicators.

After producing this consolidated set of indicators, a complementary analysis of the arguments produced in the nominal group sessions was performed. The aim was to carry out a thematic analysis of the arguments for and against the proposed indicators.¹⁷

Next, the approved indicators were displayed, by theme, in a matrix that allowed for comparison of the indicators produced by the research team and by the experts. This enabled repetitions to be identified and some proposals to be grouped together. The wording was also reviewed and standardized, finally producing a total list of 113 indicators.

Stage 4 involved the selection of experts in all regions of Brazil, followed by submission of the set

of indicators to this group, using the Delphi Method.

In the present study, each expert received a questionnaire containing instructions for its completion, followed by the list of 113 indicators. In order to verify the consistency of the indicators, the experts were asked to assign a score of 0 to 6 on each of the following criteria: clarity of wording; relevance; ease of access to data. The experts were also asked to identify what they considered the most relevant indicator in each line of action.

Analysis of consensus obtained on first consultation

The preliminary step of the analysis was to determine the values that reproduced the experts' joint opinion (by way of measurement) and the levels of consensus obtained by in this way. The cutoff point between high and low consensus could be defined by the expert group itself or using statistical methods.⁸

The experts' opinion was analyzed by judging the previously mentioned criteria: a) clarity of wording; b) relevance; and c) ease of access to data.

We used the median (M_d), which is considered a good approximation of how representative each criterion is, in combination with the interquartile range (d_q), an approximation of the degree of consensus.¹⁸ The median is defined as a measure of the central position, while the interquartile range measures data dispersion, expressed as $d_q = q_3 - q_1$, where q_1 and q_3 represent the first and third quartile, respectively. In other words, a search was made for the scores assigned to the indicators, in addition to analyzing the degree of consensus observed.

One characteristic of these measures (M_d and d_q) is the fact that they are not influenced by the occurrence of extreme values, as would be the case with the mean or standard deviation.¹⁹ Another reason for using the median rather the mean is the "arbitrary nature" of ordinal measures, where the values comprising the scale do not have a logical attribute or nature, but are the result of personal judgments.

The aim was thus to determine which values for these previously defined medians showed that the indicator achieved a good score (according to the group's opinion) and that the values reflected a group consensus.

The indicator's score was considered high if it achieved an M_d of between 4 and 6, medium or fair if it was 3, and low if the M_d was 2 or less.

A $d_q \leq 1$ suggested that the median showed a high degree of consensus, while $d_q \geq 2$ suggested a concentration around more than one position, without approaching a consensus. Consensus was

defined as values with $d_q \leq 1$, in which case they did not need to be submitted to new expert judgment.

Thus, according to our definition, indicators that were highly representative of the clarity, relevance, ease of access, from 4 to 6, and high consensus on all the criteria were immediately approved and did not have to undergo a new assessment by the experts. Indicators were also approved during the first consultation, even if they had not reached consensus for the three variables, but if they only varied from 4 to 6. These were not resubmitted, because the lack of consensus ranged among very positive values.

Indicators were discarded if they received $M_d=3$ (fair approval) for any variable plus a high degree of consensus or $M_d < 3$ (low approval), regardless of whether they achieved consensus.

In any other scenario, the indicator was resubmitted to the experts.

Based on these criteria, of the 113 indicators submitted to the experts in the first consultation, 23 were approved, 17 were rejected, and 73 were submitted to a second consultation.

Second expert consultation (Delphi method). Final analysis

The 73 indicators that failed to obtain initial consensus were resubmitted for scrutiny by the 164 experts participating in the study. After various strategies seeking produce further feedback on questionnaires in the second round, 120 experts responded (73.2%). During this second consultation, only the *access to data* variable was submitted (*the clarity and relevance* variables having obtained good-to-excellent scores for the median and low interquartile ranges).

As recommended for the Delphi Method,^{6,20} experts received the information on their previous response and the group's consolidated response, and were asked whether they wished to change or maintain their response.

Given the experts' heterogeneity (coming from various sectors), we felt it was appropriate to ask them about their current degree of certainty (on a scale from 0 to 100%) concerning access to the data included in each indicator. For example, a health expert might be certain about the data from his or her area of work, but might tend to guess or mistakenly imagine that data on the Tutelary Councils were nonexistent. The purpose of this measure was to minimize the impact of responses on *the access to data* variable for the indicator (a strategically relevant question) that were clearly based on common

sense. During this second consultation, measures of the indicator's consistency (median and interquartile range) were only calculated for responses whose *degree of certainty* was 60% or greater, a percentage established by the group of experts. The N value thus varied for each response.

In this second Stage of analysis, the same exclusion criteria were adopted as in the first phase (high consensus concerning low representativeness, $M_d \leq 3$, for all the criteria), as were the criteria for non-consensus (interquartile ranges >1).

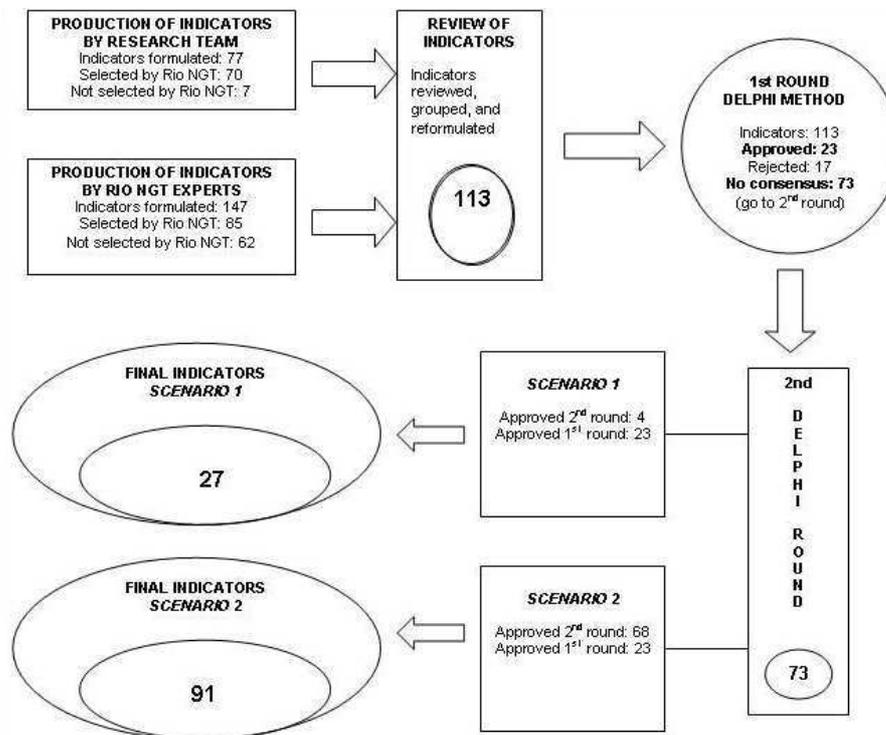
Based on these procedures, total approval was obtained (in the two Delphi rounds) for 27 indicators, while 86 indicators were rejected.

Based on the most flexible cutoff point (median greater than or equal to 3 and interquartile range >1), accepting indicators with "medium ease in data access", a completely different picture emerged: only 22 indicators were rejected, while 91 were approved. Both scenarios consist of offers of "menus" for the evaluation team. The first scenario is a more concise group of indicators, with high consensus, while the second presents a more extended group of indicators, but with less power of consensus.

Figure 1 summarizes the elaboration and selection of indicators in this study.

Figure 1

Flowchart for construction and selection of indicators for the 5 thematic lines.



Discussion

The experience reported here used two well-known methods for producing consensus. These techniques allowed validation of the indicators proposed by the team as well as the production of new indicators. The systematic analysis of guidelines from Brazil's

prevailing policies was enhanced by the professional practice and extensive day-to-day experience of experts in the various fields that work with the issue.

One parameter for these indicators was the use of construct and content validity. Construct validity deals with the relationship between theoretical concepts regarding the issue under study and their

operationalization in measures. In other words, it deals with the consistency with which indicators measure the explanatory connections supported by a given theory.²¹ Thus, as the literature confirms, consensus-generating techniques (Nominal Group and Delphi,) have an impact on the indicators' construct validity,^{12,22} since they allow criticisms to be made of the translation of the theoretical frameworks contained in the respective documents (public policies taken as the reference for actions to be implemented by the municipalities) and the proposed indicators. The indicators thus also gained content validity,²³ since the experts created indicators and critically analyzed those elaborated by the research team, in an exercise that evaluated the indicators' capacity to measure all the dimensions contained in each of the five areas of evaluation mentioned.

One innovation in this study was the joint use of NGT and the Delphi method. The output generated by the NGT became a preliminary stage for Delphi, dispensing with the initial phase of brainstorming and generation of topics. This combination allowed the elaboration of new indicators based on a face-to-face debate, by means of a dialogical and potentially richer process as compared to consulting the experts by internet.

The literature includes various discussions of the issue of the adoption of experts and the definitions used to identify them. Criticism ranges from the use of non-statistical samples to the criteria for defining the number of participants and their expertise, issues on which there is no consensus among authors.^{20,24} Since there is no information on the number of Brazilian experts on violence against children and adolescents in the various sectors (education, health, law enforcement, etc.), convenient samples constitute the most feasible option.

Heterogeneity in the sample of experts is considered a favorable strategy for increasing the validity.²⁵ The current study thus sought to ensure the heterogeneity of participants (by regional origin and sector). The concept adopted for expertise was quite inclusive and also prioritized the diversity of professional experience (managers, providers of different forms of care, and academics). Still, this

flexibility demanded a consultation regarding the "degree of certainty" in the answers provided by these experts, used as a resource for more reliable approximation to the responses.

As for the pattern of participation in the consultation, authors of Delphi studies acknowledge low adherence as a common problem^{6,26} Still, with regard to the scope of the consultation, Delphi studies operate with sample sizes that may vary considerably. Based on the methodological choices or scarcity of experts in a given field, some Delphi studies are run with 10 experts or even fewer.²⁷ Some authors¹¹ contend that 15 to 30 experts are sufficient to employ the method. It is thus a highly positive indication that the current study used 164 experts from all regions of Brazil.

Despite the long amount of time required for the consultations and the need to insist that the experts return the questionnaires, it proved possible to include a diverse, broad set of experts from the entire country, which would have been impossible using face-to-face consultations.

Finally, as some authors have pointed out,²⁰ the fact that responses have been obtained by consensus-generating techniques does not guarantee that the correct response has been found. The indicators selected using these two consensus techniques are not always (or not necessarily) the best indicators available. They reflect one given level of technical knowledge and social awareness concerning the issue of violence against children and adolescents as expressed by the experts consulted. They nevertheless constitute a wealth of knowledge with special legitimacy, owing to their collective and participatory production, anchored in both professional expertise and broadly acknowledged parameters.

The creation of two sets of indicators, one more restricted, but with greater consensus and another with broader indicators, although lower levels of consensus, expands the number of evaluation options. Considering that there is little experience of evaluation in the field of action against violence, this perspective seems justifiable, as it expands the number of choices available.

References

1. Sistema de Informação para a Infância e Adolescência. Módulo 1. Secretaria Especial dos Direitos Humanos. Brasil. [Accessed 19/3/2010]. Available in: www.mj.gov.br/sipia/
2. Secretaria Especial dos Direitos Humanos. Disque 100: cem mil denúncias e um retrato da violência sexual infantil-juvenil. Programa Nacional de Enfrentamento da Violência Sexual contra Crianças e Adolescentes. Secretaria Especial dos Direitos Humanos. Brasil. [Accessed 19/03/2010]. Available in: <http://www.bancadigital.com.br/sedh/reader2/?pID=4>.

3. Secretaria Especial dos Direitos Humanos. Matriz Intersecretorial de Enfrentamento da Exploração Sexual Comercial de Crianças e Adolescentes. Programa Nacional de Enfrentamento da Violência Sexual contra Crianças e Adolescentes. Secretaria Especial dos Direitos Humanos. Brasil. [Accessed 27/5/2008]. Available in: http://www.caminhos.ufms.br/matrizedados/mapa_matriz.html
4. Donabedian AMD. Criteria and standards for quality assessment and monitoring. *QRB*. 1986; 12: 99-108.
5. Medina MG, Silva GAP, Aquino R, Hartz ZMA. Uso de modelos teóricos na avaliação em saúde: aspectos conceituais e operacionais. In: Hartz ZMA; Vieira-da-Siva LM, editor. Avaliação em Saúde: dos modelos teóricos à prática na avaliação de programas e sistemas de saúde. Salvador / Rio de Janeiro: Fiocruz; 2005. p. 41-63.
6. Wright JTC, Giovinazzo RA. Delphi: uma ferramenta de apoio ao planejamento prospectivo. *Caderno de Pesquisas em Administração*. São Paulo; 2000. v.01, n.12, 2º trim.
7. EuropeAid. Joint Evaluation Unit. Guidelines for geographic and thematic evaluations. [Accessed 26/3/2008]. Available in: http://ec.europa.eu/europeaid/evaluation/methodology/methods/mth_pps_en.htm
8. Jones J, Hunter D. Qualitative Research: Consensus methods for medical and health services research. *British Med J*. 1995; 311: 376-80.
9. Nagahama EEI. Avaliação da implantação de serviços de saúde reprodutiva no Município de Maringá, Paraná, Brasil. *Cad Saúde Pública*. 2009; 25: 279-90.
10. Piola SF, Vianna SM, Vivas-Consuelo D. Estudo Delphi: atores sociais e tendências do sistema de saúde brasileiro. *Cad Saúde Pública*. 2002; 18: 181-90.
11. Wright JTC, Spers RG. O país no futuro: aspectos metodológicos e cenários. *Estudos Avançados*. 2006; 20: 13-28.
12. Banco Interamericano de Desenvolvimento. Matriz de marco lógico. Uma ferramenta de elaboração de projetos. [Accessed 10/01/2008]. Available in: www.enap.gov.br/downloads/ec43ea4fAvaliacao_pratica_marco_logico.pdf
13. Duggan EW, Thachenkary CS. Integrating nominal group technique and joint application development for improved systems requirements determinations. *Inf Manage*. 2004; 41: 399-411.
14. Friedman VJ, Rothman J, Withers B. The power of why: engaging the goal paradox in program evaluation. *Am J Eval*. 2006; 27: 201-18.
15. Perry J, Linsley S. The use of the nominal group technique as an evaluative tool in the teaching and summative assessment of the inter-personal skills of student mental health nurses. *Nurse Educ Today*. 2006; 26: 346-53.
16. Mattar FN. Pesquisa de marketing. Edição Compacta. São Paulo: Atlas; 2001.
17. Bardin L. Análise de conteúdo. Lisboa: Editorial Presença; 1970.
18. Bernal-Delgado E, Peiró S, Sotoca R. Prioridades de investigación en servicios sanitarios en el Sistema Nacional de Salud. Una aproximación por consenso de expertos. *Gac Sanit*. 2006; 20: 287-94.
19. Bussab WO, Morettin PA. Estatística Básica. 4 ed. São Paulo: Atual; 1987.
20. Keeney S, Hasson F, Mckenna HA. Critical review of the Delphi technique as a research methodology for nursing. *Inter J Nurs Stud*. 2003; 38: 195-200.
21. Deslandes SF, Assis SG. Abordagens quantitativa e qualitativa em saúde: o diálogo das diferenças. In: Minayo MCS, Deslandes SF, editores. Caminhos do pensamento: epistemologia e método. Rio de Janeiro: Fiocruz; 2003. p. 195-223.
22. Cosendey MAE, Hartz ZMA, Bermudez JAZ. Validation of a Tool for Assessing The Quality of Pharmaceutical Services. *Cad Saúde Pública*. 2003; 19: 395-406.
23. Reichenheim M, Moraes CL. Buscando a qualidade das informações em pesquisas epidemiológicas. In: Minayo MC, Deslandes SF, editores. Caminhos do pensamento: epistemologia e método. Rio de Janeiro: Editora Fiocruz; 2002. p. 227-54.
24. Baker J, Lovell K, Harris N. How expert are the experts? An exploration of the concept of 'expert' within Delphi panel techniques. *Nurs Res*. 2006; 14: 59-70.
25. Mead D, Moseley L. The use of Delphi as a research approach. *Nurs Res*. 2001; 8: 4-23.
26. Keeney S, Hasson F, Mckenna HA. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J Adv Nurs*. 2006; 53: 205-12.
27. Noblat L, Oliveira MGG, Santos R, Noblat ACB, Badaró R. Validation of criteria for nosocomial use of amikacin in Brazil with the Delphi technique. *Braz J Infect Dis*. 2006; 10: 173-8.

Recebido em 31 de março de 2010

Versão final apresentada em 9 de agosto de 2010

Aprovado em 1 de setembro de 2010