

Avaliação do impacto de programas nutricionais

Evaluation of nutritional programs's impact

Iná da Silva dos SANTOS¹

RESUMO

Muito se tem deixado de aprender sobre programas e intervenções que realmente funcionam, em decorrência tanto da escassez quanto da variada qualidade metodológica das avaliações de impacto. Avaliações de impacto são estudos que medem o impacto sobre a saúde diretamente atribuível a uma política pública ou programa específico, independentemente do efeito de outros potenciais fatores explanatórios. Com o objetivo de discutir, à luz da epidemiologia, os tipos de estudos que podem ser empregados para a avaliação de impacto de programas em nutrição, são apresentados os delineamentos conforme o grau de inferência necessária para a avaliação (adequação, plausibilidade e probabilidade). São apresentados exemplos de estudos de avaliação de programas com desenhos transversais, quase-experimentos e ensaios randomizados controlados. São descritas as vantagens dos desenhos observacionais sobre as avaliações conduzidas com ensaios randomizados, quando o objetivo é avaliar a efetividade dos programas em condições reais de campo. É urgente que avaliações de impacto sejam conduzidas para fornecerem subsídios aos tomadores de decisões sobre os programas que valem a pena receber investimentos ou que necessitam ser reformulados.

Termos de indexação: Alimentação. Avaliação de programas e projetos de saúde. Estudos de avaliação como assunto. Nutrição em saúde. Políticas públicas.

ABSTRACT

Persistent shortcomings in our knowledge of the actual effects of programs and interventions are due to a gap in both the quantity and methodological quality of impact evaluations. Impact evaluations are studies that measure the impact on health directly attributable to a specific program or policy, regardless of other potential explanatory factors. The aim of this manuscript is to present the types of epidemiological studies that can be employed in impact evaluations of nutrition programs. Study designs are presented according to the level of inference needed for the evaluation (adequacy, plausibility and probability). Examples of program evaluations using cross-sectional, quasi-experimental, and randomized controlled trials are discussed. The advantage of the observational studies over randomized controlled trials when the aim is to assess programs effectiveness in real world conditions is highlighted. Impact evaluations are urgently needed to subsidize decision makers on what programs should continue to receive investments and what programs need to be reformulated.

Indexing terms: Feeding. Program evaluation. Evaluation studies as topic. Nutrition public health. Public policies.

¹ Universidade Federal de Pelotas, Centro de Pesquisas Epidemiológicas, Programa de Pós-Graduação em Epidemiologia. R. Marechal Deodoro, 1160, 3º piso, 96020-220, Pelotas, RS, Brasil. E-mail: <inasantos@uol.com.br>.

INTRODUÇÃO

Muito se sabe atualmente sobre a prevalência e as causas dos problemas nutricionais que afligem as populações dos países em desenvolvimento. Sabe-se também que é necessária a combinação de uma tríade - investimentos, vontade política e implementação de políticas públicas efetivas - para que esses problemas possam ser enfrentados¹. Em geral, as intervenções em larga escala, que visam a prevenção ou tratamento desses problemas, são implementadas sem a correspondente realização de avaliações de seu impacto sobre indicadores do estado nutricional ou de saúde. O conhecimento sobre quais políticas públicas são efetivas, no entanto, depende da condução de avaliações deste tipo. Estes estudos medem o impacto diretamente atribuível a uma política pública ou programa específico, independente do efeito de outros potenciais fatores explanatórios. Apesar da importante necessidade de saber o que funciona ou não, as avaliações de impacto são raras e a qualidade metodológica das que são realizadas é muito variável. A falta de informações sobre o que funciona ou não deixa os tomadores de decisão com pouca base para defender a razão de seus investimentos ou, se necessário, para melhorar os programas².

A nutrição, particularmente a nutrição infantil, tem sido alvo de inúmeros programas, que visam a prevenção da desnutrição e das carências nutricionais específicas³⁻⁶. Somas vultosas de recursos têm sido gastas por governos e fundações internacionais e muito pouco se sabe sobre o real efeito desses programas ou intervenções sobre as deficiências que pretendem prevenir ou tratar. As poucas avaliações realizadas mostram, em geral, efeito nulo ou apenas discreto. Porém, além dos múltiplos fatores envolvidos na causalidade da desnutrição e das deficiências nutricionais, capazes de sobrepujar o efeito real dos programas, outros, ligados à natureza da avaliação, podem contribuir para essa série de fracassos⁷.

O objetivo desse artigo é discutir, à luz da epidemiologia, os tipos de estudos que podem

ser utilizados para a avaliação de impacto de programas em nutrição. Nesse contexto, "impacto" significa o efeito de uma intervenção ou programa sobre sua população alvo. Especificamente, o trabalho propõe-se a apresentar os delineamentos epidemiológicos que podem ser empregados para avaliações de impacto de programas ou intervenções de larga escala.

Dois aspectos são importantes quando se planeja e avalia um programa de intervenção para melhorar a nutrição infantil: a) o tipo de alimento, suplemento e/ou aconselhamento nutricional que é dispensado; e b) o processo empregado para que o alimento, suplemento ou as recomendações atinjam os beneficiários. Um mecanismo perfeito de distribuição terá o impacto positivo esperado sobre o estado nutricional da criança somente se os alimentos ou suplementos oferecidos forem nutricionalmente adequados ou se o aconselhamento nutricional for bem concebido. Por outro lado, suplementos ou recomendações nutricionais apropriados serão inadequados se forem incompatíveis com a cultura local, se forem economicamente inacessíveis para a população, se houver entraves estruturais ou de processo na distribuição ou se os beneficiários não estiverem suficientemente motivados para aceitá-los⁸.

Independente do tipo de estudo que se deseje empregar para avaliação de impacto, é muito importante identificar o modelo lógico (a cadeia causal) que reflita os mecanismos por meio dos quais o programa, se efetivo, deverá funcionar. Como exemplo, a Figura 1 mostra, esquematicamente, o modelo de impacto da Atenção Integrada às Doenças Prevalentes na Infância (AIDPI)⁹, discriminando os três pilares sobre os quais se apóia a estratégia de redução da mortalidade infantil e melhora do estado nutricional. Para que o efeito da intervenção possa ser verificado, é necessário que o processo nessas três instâncias seja qualificado. Primeiro, deve haver melhorias nos sistemas de saúde, mediante a disponibilidade de vacinas, medicamentos, supervisão da qualidade da assistência etc.; segundo, os profissionais de saúde devem ser treinados na estratégia e devem

ser realizadas visitas de seguimento para garantir a sustentabilidade; e terceiro, intervenções específicas devem ser direcionadas às famílias e comunidades. De acordo com esse modelo, as melhorias nos sistemas de saúde e o treinamento dos profissionais levariam a um aumento na qualidade da atenção nos serviços de saúde e à conseqüente melhora no manejo domiciliar de casos e adesão às recomendações. As intervenções direcionadas às famílias e às comunidades, por sua vez, resultariam em melhora na busca por cuidados de saúde e aumento na utilização de serviços de saúde. As melhorias na qualidade da atenção, no manejo domiciliar das doenças e o aumento da utilização levariam ao desfecho desejado de redução da mortalidade infantil e melhora do estado nutricional.

A construção do modelo lógico é fundamental para definir os indicadores que poderão ser utilizados na avaliação, bem como auxiliar na identificação dos potenciais confundidores do efeito do programa sobre o estado nutricional. Entende-se por “confundidores” fatores outros que não o próprio programa que concorram para o

alcanço do mesmo efeito sobre a saúde. Esses fatores incluem características contextuais (políticas, sociais, econômicas, agrícolas etc.) e a presença de outros investimentos ou programas que podem impactar o mesmo problema para o qual o programa se dirige. Voltaremos à idéia de confundidores mais adiante nesse manuscrito.

Tipos de indicadores a utilizar

Um modelo de avaliação de programas e projetos em saúde pública propõe que a avaliação seja realizada em um crescendo, que inicia pela medição de indicadores de oferta e se estendendo para indicadores de utilização e cobertura, antes que os de impacto sejam medidos⁹. De acordo com esse modelo (Figura 2), é investigado, inicialmente, se as intervenções e suas estratégias de implementação são tecnicamente bem fundamentadas e apropriadas ao contexto epidemiológico e assistencial do país (avaliação das políticas e do planejamento baseado em resultados). Logo, se cuidados adequados estão sendo oferecidos nos

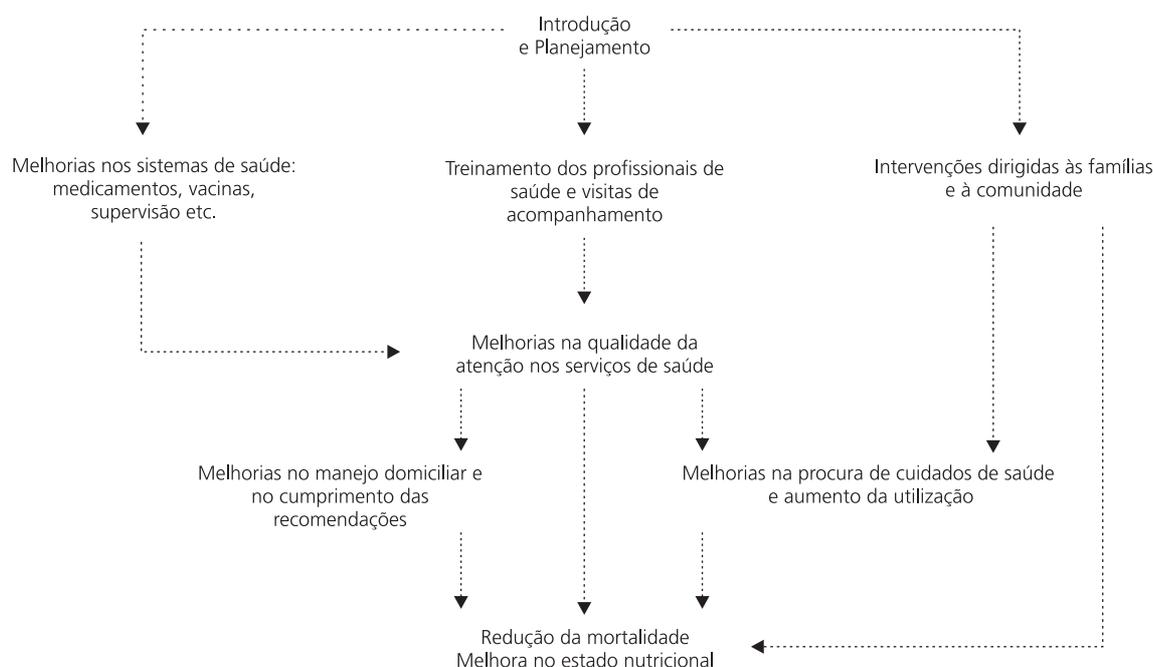


Figura 1. Modelo do impacto da Atenção Integrada às Doenças Prevalentes na Infância²².

serviços de saúde e na comunidade (indicadores de oferta). Depois, se esses cuidados estão sendo usados pela comunidade (indicadores de utilização), de forma a alcançar um nível adequado de cobertura efetiva na população (indicadores de cobertura). Somente depois de respondidas afirmativamente essas perguntas é que as avaliações de impacto sobre a morbi-mortalidade, o estado nutricional ou outros indicadores biológicos ou do comportamento deveriam ser realizadas (indicadores de impacto).

A adequada documentação de indicadores de processo (oferta, utilização e cobertura), durante a implementação do programa, é extremamente importante para entender os fatores que intervieram no impacto. No caso de programas que não logram alcançar impacto, por outro lado, a avaliação do processo permite identificar as medidas que devem ser introduzidas para melhorar o programa.

Tipos de estudos para avaliar impacto de programas de nutrição

Muitas avaliações de impacto são realizadas por intermédio de estudos de eficácia, que investigam se os programas funcionam em condições ideais. Nessas avaliações os programas, em geral, são implementados de forma integral, com cuidadosas estratégias para que alcancem uma população alvo com grande necessidade de intervenção, garantindo alta cobertura e qualidade. Amparados em estritos critérios de seleção e no princípio da randomização (que pareia os grupos randomizados quanto ao prognóstico), raramente essas avaliações descrevem os fatores contextuais (essenciais para decidir sobre a capacidade de generalização dos resultados) presentes nos locais e serviços onde se realizam as avaliações.

Os estudos de efetividade, por sua vez, são realizados para avaliar se o programa funciona

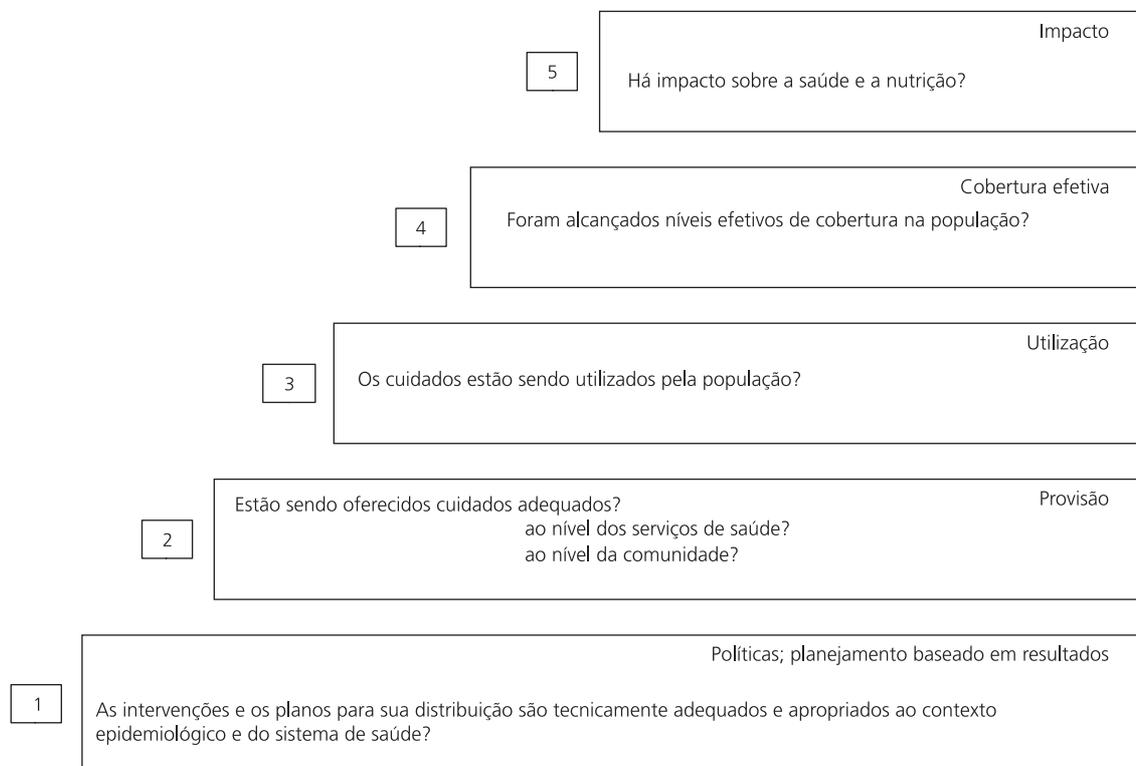


Figura 2. Modelo para avaliação de programas em saúde pública²².

em condições da vida real^{9,10}. Nesse caso, os programas são implementados de forma imperfeita, com qualidade intermediária, alcançam cobertura parcial e, em geral beneficiam os menos pobres, que por não necessitarem tanto da intervenção, respondem com menor impacto do que nos estudos de eficácia. É também possível que nas avaliações de efetividade estejam ausentes os fatores contextuais que estavam presentes nos locais onde se realizaram as avaliações de eficácia (essenciais para o êxito do programa)

O modelo de avaliação de Habicht *et al.*¹⁰ propõe um segundo eixo (além do eixo dos indicadores de oferta, utilização, cobertura e impacto) com três níveis de inferência para a avaliação: adequação, plausibilidade e probabilidade. O nível de inferência depende do tipo de estudo empregado para a avaliação. A decisão quanto ao tipo de estudo a empregar, por sua vez, depende do objetivo da avaliação e do grau de certeza que se necessita ter de que o resultado observado sobre o estado nutricional da população alvo é decorrente do programa propriamente dito. Um detalhado manual sobre como planejar a implementar avaliações de programas, inclusive avaliações de impacto, pode ser encontrado em outra fonte¹³.

Tipos de estudos para avaliações de adequação

Se o objetivo da avaliação for verificar se o programa está atingindo os objetivos a que se propôs (por exemplo, reduzir em 20% o déficit de peso para idade entre crianças em risco de desnutrição, reduzir a prevalência de anemia para 15% etc.) ou se as tendências nos indicadores de impacto ocorreram na direção esperada e com magnitude adequada, estudos descritivos que meçam a prevalência do problema que se quer corrigir serão suficientes.

Essas avaliações poderão ser realizadas em apenas um ou em mais de um tempo. Incluem-se nesse grupo os estudos transversais, realizados em uma única ocasião, durante ou ao final do programa; estudos tipo antes-e-depois, com uma medida

de linha de base (antes da implementação do programas); e as séries temporais, que envolvem medidas repetidas para identificação de tendências nos indicadores estudados.

Um programa de redução da mortalidade infantil na cidade de Pelotas, no sul do Brasil, iniciado em 2005, por exemplo, tinha como meta reduzir a mortalidade infantil de 20 para 15 por mil nascidos-vivos, o que correspondia a menos de 65 mortes em menores de um ano, por ano. Ao final do primeiro ano de intervenção, a mortalidade havia caído para 15,5 por mil nascidos-vivos e haviam ocorrido 66 mortes de menores de um ano residentes no município (Figura 3). O resultado da avaliação de adequação, portanto, mostrou um sucesso parcial.

O estudo de base populacional de Assunção *et al.*¹⁴, que mediu a prevalência de anemia entre crianças de até seis anos de idade, antes e depois que o cumprimento da lei que obriga a fortificação das farinhas de trigo e milho com sais de ferro se tornasse obrigatória no Brasil, é outro exemplo de uma avaliação de adequação com séries temporais. Por determinação do Ministério da Saúde, toda a farinha de trigo e milho produzida no país a partir de julho de 2004 deveria receber adição de ferro. Os compostos de ferro de grau alimentício a serem utilizados - sulfato ferroso desidratado, fumarato ferroso, ferro reduzido e ferro elemento - 325 *mesh Tyler*, etileno-diaminotetraacetato (EDTA) de ferro e sódio e ferro bisglicina quelado, assim como sua quantidade (4,2mg de ferro/100g de farinha), foram determinados por portaria específica¹⁵. Foram realizados três inquéritos transversais na cidade de Pelotas (RS). Entre maio e junho de 2004, anteriormente à obrigatoriedade de fortificação das farinhas, os níveis de hemoglobina foram investigados em um grupo de crianças de zero a cinco anos de idade (estudo de linha de base). Posteriormente, transcorridos doze e 24 meses de vigência da lei (respectivamente, em 2005 e 2006), os níveis de hemoglobina foram medidos em outros grupos de crianças, com idade e nível socioeconômico similares aos do grupo de linha

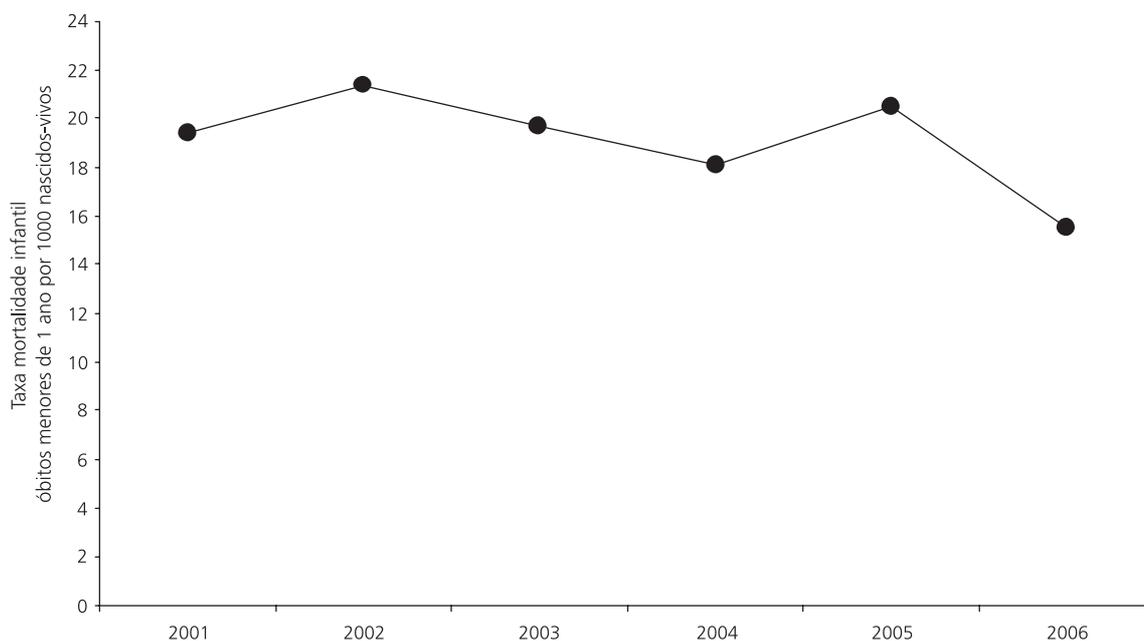


Figura 3. Taxa de mortalidade infantil por ano. Pelotas (RS), 2007.

de base. Nesse estudo, não foi examinado o mesmo grupo de crianças nos três momentos porque, por razões éticas, os casos de anemia diagnosticados eram encaminhados para tratamento.

As avaliações de adequação normalmente são fáceis de ser conduzidas, não requerendo treinamento específico em pesquisa. Principalmente quando negativas (ou seja, quando mostram que não houve impacto), são muito úteis para repensar o programa. Quando mostram impacto positivo sobre os indicadores, por outro lado, os métodos descritos a seguir podem ser úteis para avaliar se esse resultado é de fato devido ao programa ou não.

Estudos para avaliações de plausibilidade

As avaliações de plausibilidade respondem se é provável que o impacto observado tenha decorrido do programa. Esse tipo de avaliação requer que sejam descartados os efeitos de fatores externos ao programa, havendo, portanto, a necessidade de um grupo de comparação (grupo

controle) para ajuste de fatores de confusão. Como a definição do grupo que recebe o programa (grupo inter-venção) ou não (grupo controle) não é aleatória, esses estudos são conhecidos como "quase-experimentos".

Um estudo não controlado, realizado no Rio de Janeiro, havia detectado impacto positivo do Programa Nacional de Leite¹⁶ na recuperação nutricional de crianças desnutridas¹⁷. Uma avaliação de plausibilidade do impacto do mesmo programa foi, posteriormente, realizada no estado de Alagoas e está esquematizada na Figura 4. Foram selecionadas crianças recém-ingressadas no programa em dez municípios com alta cobertura do programa (grupo intervenção) e um número equivalente de crianças com indicador peso/idade abaixo do percentil 10 de referência (elegíveis, portanto para receberem o suplemento) de dez municípios com baixa cobertura e que não conseguiram vaga no programa (grupo controle)¹⁸. As crianças foram avaliadas antropometricamente (peso, comprimento e composição corporal com isótopos) em duas ocasiões: ao ingressarem no estudo e seis meses mais tarde. A avaliação inicial mostrou que havia diferença estatisticamente

significativa entre os dois grupos quanto a características sociais: as crianças do grupo que recebia o programa pertenciam a famílias mais pobres e os pais tinham menor escolaridade. Após seis meses, ajustando para essas variáveis, a avaliação mostrou que não havia diferença entre os dois grupos quanto ao estado nutricional. Além dos indicadores de impacto, foram coletadas informações sobre o processo (redistribuição do suplemento dentro da família, regularidade na entrega do suplemento nos serviços de saúde etc.). Ficou claro que o suplemento era consumido não apenas pela criança beneficiária do programa, mas também por outros membros da família, e que a distribuição do suplemento não era feita regularmente nos postos de saúde, inclusive durante o período da avaliação. A deficiência neste indicador de oferta, acrescida da utilização parcial do suplemento, provavelmente foram os responsáveis pelo insucesso do programa.

Em decorrência da não-randomização, a desejável "igualdade" entre os grupos quanto à distribuição de características confundidoras não pode ser garantida, sendo que as crianças do grupo controle, embora pobres e desnutridas, pertenciam a famílias com condições sociais levemente melhores (mas com diferença estatisticamente

significativa) do que as crianças beneficiadas pelo programa. Tal fato implicou na necessidade de se descartar (à parte a possibilidade de real não efetividade do programa) explicações alternativas para os resultados encontrados. As análises ajustadas realizadas com esse fim, no entanto, não mudaram o resultado observado.

Uma avaliação do impacto do programa Bolsa Família também utilizou um delineamento quase-experimental¹⁹. Com o objetivo de determinar e estimar a magnitude dos resultados devidos exclusivamente ao Programa Bolsa Família, foram analisados dois momentos após a inclusão da família no programa: o primeiro, após seis meses de pagamentos e, o segundo, com dois anos. Em relação às crianças do grupo controle, tanto a frequência quanto a gravidade dos problemas nutricionais eram maiores no grupo beneficiário ao ingressarem no programa. Como mostrou o resultado de um índice composto por essas duas variáveis (gravidade e frequência de problemas nutricionais), essa diferença entre os dois grupos foi diminuindo gradualmente ao longo dos primeiros 15 meses do programa, em decorrência da melhora do estado nutricional do grupo beneficiário (intervenção).

Avaliação do impacto do Programa Nacional do Leite no estado de Alagoas

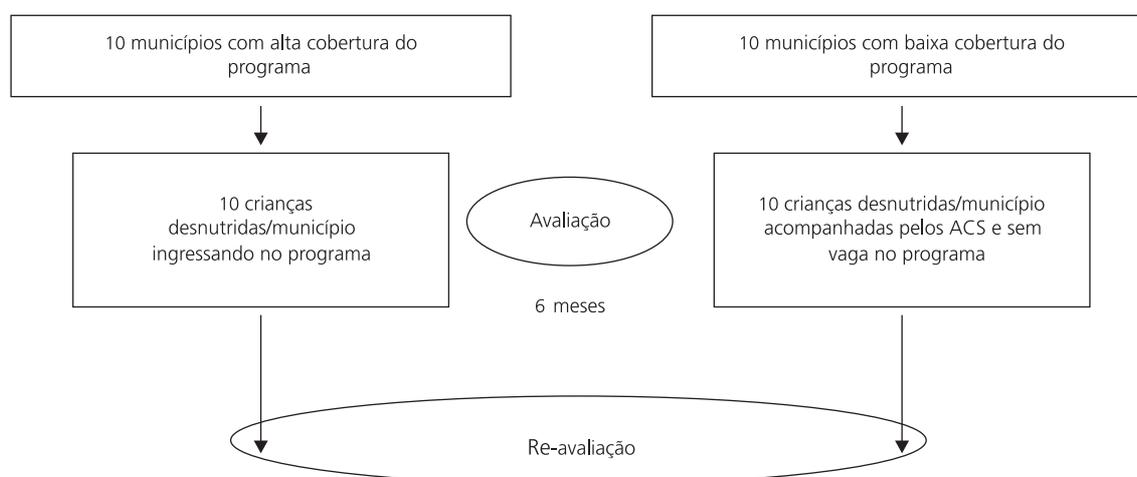


Figura 4. Estrutura esquemática de um estudo quase-experimental utilizado na avaliação do programa de leite no estado de Alagoas¹⁷.

Os estudos de casos e controles têm sido muito pouco utilizados na avaliação de programas. Hipoteticamente, a estrutura de um estudo de casos e controles para avaliar o impacto de um programa de suplementação de vitamina A, por exemplo, poderia ser organizado como mostra a Figura 5. Crianças identificadas com deficiência seriam arroladas como casos. Os controles (emparelhados ou não) seriam selecionados das mesmas comunidades de onde emergiram os casos. Se o programa tivesse impacto positivo, seria esperado que o *odds* de exposição entre os casos (casos expostos ao programa: casos não expostos ao programa) fosse menor que o *odds* de exposição entre os controles. Dada sua natureza retrospectiva, os estudos de casos e controles, em geral, são conduzidos mais rapidamente do que os quase-experimentais e deveriam ser empregados com maior frequência na avaliação de programas.

Estudos para avaliações de probabilidade

As avaliações de probabilidade, conduzidas exclusivamente através de ensaios randomizados,

priorizam a validade interna da avaliação: a randomização reduz o risco de viés de seleção e de confusão e o mascaramento reduz o risco de viés de informação. Estes estudos são essenciais para determinar a eficácia de novos programas, sendo particularmente adequados para verificar o efeito de intervenções cuja cadeia causal entre a intervenção e o resultado é curto, como ocorre com o efeito biológico de medicamentos, vacinas, suplementos nutricionais etc. Além disso, esses estudos geralmente medem eficácia e não efetividade, mesmo quando não são tomadas medidas específicas para adesão ao protocolo do programa (ensaios de efetividade), pois a própria presença de um grupo de avaliadores pode influenciar o desempenho do programa, em decorrência do efeito Hawthorne²⁰.

A cadeia causal de um ensaio comunitário randomizado de educação nutricional, por exemplo, requeria o cumprimento de pelo menos sete passos, antes de obter a melhora nutricional em crianças menores de dois anos de idade: implementação nacional do programa, treinamento dos profissionais de saúde, aumento do conhecimento dos profissionais de saúde em educação nutricional, melhora do desempenho dos profissionais

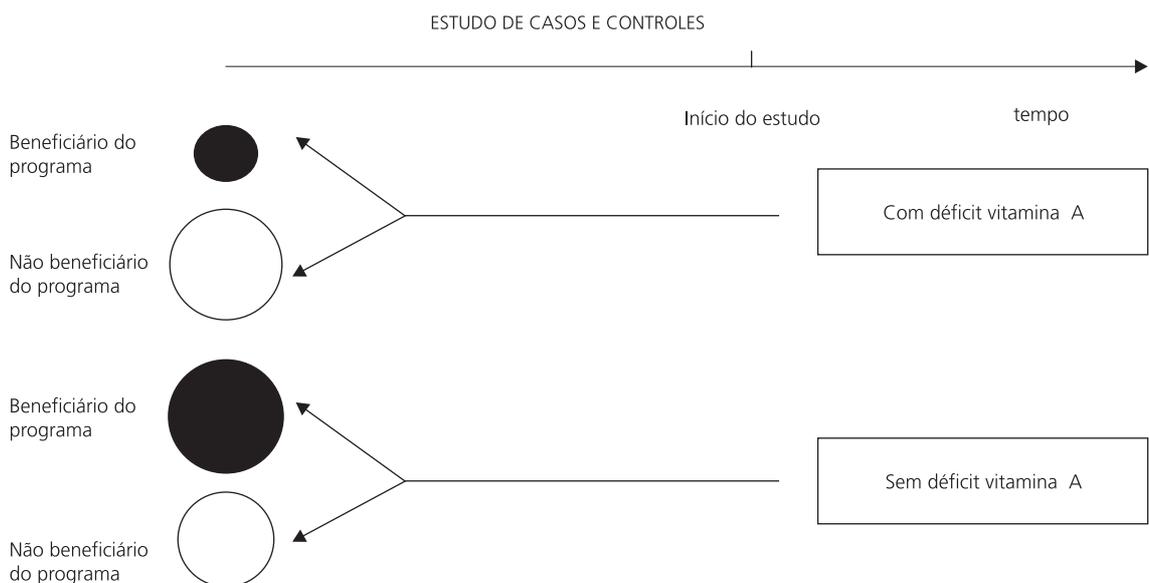


Figura 5. Modelo hipotético de um estudo de casos e controles para avaliar o impacto de um programa destinado a reduzir a deficiência de vitamina A.

de saúde em educação nutricional, aumento do conhecimento em nutrição das mães atendidas pelos profissionais treinados, melhora na alimentação das crianças e aumento na ingestão de calorias. Um ensaio comunitário, randomizado, controlado, foi implementado com o objetivo de medir a eficácia do aconselhamento nutricional dirigido a mães de menores de 18 meses de idade durante consultas médicas²¹. Os 28 postos de saúde da cidade foram emparelhados quanto a indicadores sociais e de saúde da comunidade e um posto de cada par foi sorteado para receber o treinamento (grupo intervenção). Os médicos do grupo intervenção foram treinados em aconselhamento nutricional na infância conforme as normas da estratégia AIDPI. Em seis meses, as crianças do grupo intervenção que ingressaram no estudo com 12-17,9 meses de idade ganharam em média 1,48kg, enquanto que as do grupo controle (atendidos por médicos não treinados em educação nutricional) ganharam em média 1,14kg ($p < 0,05$). Em uma larga cadeia de causalidade como a implicada nesse estudo, é necessário um alto desempenho em cada passo do processo para que um efeito dessa magnitude possa ser detectado. Dificilmente se alcança tal nível de excelência em condições reais, especialmente em países em desenvolvimento, onde os programas são implementados.

CONCLUSÃO

Foram apresentados os tipos de indicadores e delineamentos epidemiológicos que podem ser utilizados para avaliação do impacto de programas nutricionais. Há uma profunda distância entre o número de programas nutricionais implementados em todo o mundo e o que sabemos sobre sua efetividade a partir de avaliações de impacto. Essa distância deve-se à escassez das avaliações de impacto e à variada qualidade metodológica das avaliações implementadas. Muito se tem deixado de aprender sobre programas e intervenções que realmente funcionam em decorrência dessa escassez. Em uma época em que os recursos são

escassos para atender todas as necessidades sociais e de saúde das populações mais pobres, é urgente que avaliações de impacto sejam conduzidas para fornecerem subsídios aos tomadores de decisões sobre os programas nos quais vale a pena fazer investimentos e os que necessitam ser reformulados.

REFERÊNCIAS

1. Center for Global Development. When will we ever learn? Report of the Evaluation Gap Working Group. Washington (DC): 2006 [cited 2008 Mar 20]. Available from: <www.cgdev.org/section/initiatives/_active/evalgap>.
2. Bryce J, Victora CG. MCE-IMCI Technical Advisors. Ten methodological lessons from the multi-country evaluation of Integrated management of childhood illness. Health Policy Plan. 2005; 20 (Suppl 1):i94-i105.
3. World Health Organization. A critical link. Interventions for physical growth and physiological development: a review. Geneva: Who; 1999. WHO/CHS/CAH/99.3.
4. Hill Z, Kirkwood B, Edmond K. Family and community practices that promote child survival, growth and development: a review of the evidence. Geneva: WHO; 2004.
5. Allen L, Gillespie S. What works? A review of the efficacy and effectiveness of nutrition interventions. Geneva: World Health Organization; 2001.
6. Santos IS. Intervenções nutricionais na infância. In: Kac G, Sichieri R, Gigante DP, editores. Epidemiologia nutricional. Rio de Janeiro: Atheneu; 2007.
7. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ. 1996; 312:1215-18.
8. World Health Organization. Complementary feeding of young children in developing countries: a review of current scientific knowledge. Geneva: WHO; 1998. WHO/NUT/98.1.
9. Organização Pan-Americana da Saúde. Atenção Integrada às doenças prevalentes na infância (AIDPI). Washington (DC): WHO; 1999.
10. Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials. 3rd ed. New York: Springer; 1998.
11. Gordis L. Epidemiology. 3rd ed. Philadelphia: Elsevier Saunders; 2004.

12. Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Intern J Epidemiol.* 1999; 28(1):10-8.
13. Santos IS. Guia metodológico de avaliação e definição de indicadores: doenças crônicas não-transmissíveis e Rede Carmen. Brasília: Organização Pan-Americana da Saúde, 2006 [citado 2008 Fev 8]. Disponível em: <<http://www.saude.gov.br/bvs>>.
14. Assunção MCF, Santos IS, Barros AJD, Gigante DP, Victora CG. Efeito da fortificação de farinhas com ferro sobre anemia em pré-escolares, Pelotas, RS. *Rev Saúde Pública.* 2007; 41(4):539-48.
15. Agência Nacional de Vigilância Sanitária. Resolução no. 344. Revoga a resolução RDC nº 15, de 21 de fevereiro de 2000. *Diário Oficial da União* 2002; 18 de dezembro.
16. Brasil. Plano de combate à fome e à miséria: princípios, prioridades e mapa das ações de governo. Brasília: MDS; 1993.
17. Castro IRR, Monteiro CA. Avaliação do impacto do programa "Leite é Saúde" na recuperação de crianças desnutridas no Município do Rio de Janeiro. *Rev Bras Epidemiol.* 2002; 5(1):52-62.
18. Santos IS, Gigante DP, Coitinho DC, Haisma H, Valle NCJ, Valente GCM. Avaliação do impacto de um programa de suplementação alimentar para crianças desnutridas no Brasil. *Cad Saúde Pública.* 2005; 21(3):776-85.
19. Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. Avaliação do programa bolsa família: segunda fase. Brasília: Ministério da Saúde; 2005. Série C. Projetos, programas e relatórios.
20. Victora CG, Habicht JP, Bryce J. Evidence-based public health. Moving beyond randomized trials *Am J Public Health.* 2004; 94(3):400-5.
21. Santos I, Victora CG, Martines J, Gonçalves H, Gigante DP, Valle NJ, *et al.* Nutrition counseling increases weight gain among Brazilian children. *J Nutr.* 2001; 131(11):2866-73.
22. Bryce J, Victora CG, Habicht JP, Vaughan JP, Black RE. The multi-country evaluation of the integrated management of childhood illness strategy: lessons for the evaluation of public health interventions. *Am J Public Health.* 2004; 94(3):406-15.

Recebido em: 28/3/2008

Versão final reapresentada em: 7/5/2008

Aprovado em: 27/11/2008