

1. Conceitos, abordagens e estratégias para a avaliação em saúde

Ligia Maria Vieira da Silva

SciELO Books / SciELO Livros / SciELO Libros

SILVA, LMV. Conceitos, abordagens e estratégias para a avaliação em saúde. In: HARTZ, ZMA., and SILVA, LMV. orgs. *Avaliação em saúde: dos modelos teóricos à prática na avaliação de programas e sistemas de saúde* [online]. Salvador: EDUFBA; Rio de Janeiro: Editora FIOCRUZ, 2005, pp. 15-39. ISBN: 978-85-7541-516-0. Available from: doi: [10.7476/9788575415160](https://doi.org/10.7476/9788575415160). Also available from in ePUB from: <http://books.scielo.org/id/xzdnf/epub/hartz-9788575415160.epub>



All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

CONCEITOS, ABORDAGENS E ESTRATÉGIAS PARA A AVALIAÇÃO EM SAÚDE

Ligia Maria Vieira da Silva

REVISITANDO O CAMPO DA AVALIAÇÃO

A avaliação pode ser considerada como um componente das práticas presente em diversos âmbitos e campos do espaço social. As suas diversas possibilidades de expressão, nesses diferentes campos, tem gerado uma polissemia conceitual e metodológica que torna obrigatório, tanto para o gestor quanto para o pesquisador, uma explicitação das suas opções teóricas e técnicas (GRÉMY, 1983; VIEIRA-DA-SILVA e FORMIGLI, 1994; CONTANDRIOPOULOS e outros, 2000).

Entre as diversas definições existentes sobre o significado da avaliação, aquelas referentes a avaliação de programas sociais têm conseguido maior consenso. Uma delas relaciona a avaliação ao processo de determinação do esforço, mérito ou valor de algo ou do produto desse processo (SCRIVEN, 1991). Patton, ao desenvolver uma abordagem da avaliação denominada de “*avaliação focada para a utilização*” inclui na sua definição: a) a maneira de fazer a avaliação que corresponderia à coleta sistemática de informações sobre as atividades, características e resultados dos programas; b) o propósito da avaliação, ou seja, fazer julgamentos sobre os programas e/ou para subsidiar o processo de tomada de decisões sobre futuras programações (PATTON, 1997). Já Rossi e Freeman, que editam um dos mais conhecidos livros sobre avaliação, a consideram como sinônimo de pesquisa avaliativa e a definem como a aplicação sistemática dos procedimentos oriundos da investigação em ciências sociais para analisar programas de intervenção (ROSSI; FREEMAN, 2004).

Na mesma perspectiva de delimitar o campo da avaliação ao julgamento de programas sociais, Contandriopoulos formulou uma das definições mais simples e abrangentes. Segundo esse autor, a avaliação pode ser considerada como um julgamento sobre uma intervenção ou sobre qualquer dos seus componentes com o objetivo de auxiliar na tomada de decisões (CONTANDRIOPOULOS e outros, 1997). O problema com esse conceito é que a definição de intervenção feita pelo autor limita o escopo da avaliação¹ à dimensão instrumental da ação.

O conceito de campo, enquanto rede de relações entre agentes e instituições, no sentido que lhe é dado por Bourdieu (1979; BOURDIEU e WACQUANT, 1992), pode contribuir para uma concepção da avaliação muito mais abrangente do que aquela restrita à dimensão racional da ação, ou aquela que situa a avaliação em relação ao uso instrumental dos resultados (PATTON, 2001). Nessa perspectiva, a definição de avaliação proposta por Contandriopoulos será aqui ajustada ao escopo maior do campo da avaliação substituindo o conceito de “intervenção” por aquele de “práticas sociais”² – em relação às quais, as práticas de saúde podem ser consideradas como um caso particular. Isto porque, a avaliação toma por objeto desde as práticas do cotidiano até aquelas relacionadas com o trabalho, nos seus diversos âmbitos, o que envolve as intervenções sociais planejadas. O significado do que seria o “julgamento” também pode variar desde a formulação de um juízo de valor dicotômico qualitativo ou quantitativo (cuidado à saúde, bom ou ruim, cobertura vacinal igual ou superior a 70% ou inferior a 70%) até uma análise que envolva o significado do fenômeno.

A ampliação da definição de Contandriopoulos se justifica por referência à consolidação da avaliação enquanto campo. Este se revela não apenas na ampliação da produção científica específica, mas sobretudo pela rede de agentes (pesquisadores, avaliadores e gestores) e associações profissionais envolvidos (CHELIMSKY, 1997; MCDAID, 2003), pelo seu reconhecimento enquanto disciplina e profissão (CHELIMSKY e SHADISH, 1997; SMITH, 2001) bem como pela diversificação das áreas temáticas que são objeto de avaliação sistemática (SCRIVEN, 2001).

Cabe aqui uma distinção entre monitoramento e avaliação pontual. O monitoramento corresponderia ao acompanhamento sistemático sobre algumas das características dos serviços enquanto que a avaliação pontual refere-se ao julgamento sobre os mesmos em um determinado ponto do tempo.

Dessa forma, em algumas circunstâncias, o monitoramento pode produzir informações para a realização de uma avaliação mas não corresponder necessariamente a implementação da mesma. Ou seja, quando a atividade de monitoramento se resume ao registro contínuo de variáveis ao longo do tempo se superpõe à montagem de um sistema de informação. A transformação de dados em informação bem como a sua utilização para a avaliação, requer a formulação de uma pergunta específica que orientaria a análise e os possíveis sentidos a serem atribuídos aos achados empíricos. A otimização dos sistemas de informação para fins de avaliação bem como a indicação adequada acerca da necessidade ou não da realização de avaliações pontuais ou de investigações avaliativas requer certo grau de institucionalização da avaliação (HARTZ, 1999).

O ESPECTRO DA AVALIAÇÃO

Para fins didáticos, pode-se imaginar essa diversidade de possibilidades de expressão como faixas ou pontos de um espectro, cujo extremo esquerdo pode ser representado pelas avaliações presentes na vida cotidiana e o outro extremo pela investigação avaliativa³ (Figura 1). A avaliação das práticas cotidianas corresponderia a um julgamento que se faz a partir do recurso a noções oriundas do “senso comum”, a técnicas não sistemáticas de observação e à análise e formulação de juízos de valor dicotômicos e simplificados. Por exemplo, ao final de um dia de trabalho uma equipe de saúde reúne-se para “fazer um balanço” das atividades daquele dia referentes ao funcionamento de um serviço de controle de hipertensão arterial. Consideram que o balanço é “positivo” pois a demanda foi “grande” e o atendimento foi “bom” deixando os usuários “satisfeitos”. Esse balanço corresponde a um julgamento das práticas rotineiras do trabalho daquela equipe a partir da percepção dos seus componentes acerca do que seria um desempenho “positivo” no trabalho, uma “grande” demanda, um “bom” atendimento e uma adequada “satisfação” dos usuários.

Já a pesquisa avaliativa corresponderia ao julgamento que é feito sobre as práticas sociais a partir da formulação de uma pergunta não respondida ainda na literatura especializada, sobre as características dessas práticas, em geral, ou em um contexto particular, através do recurso a metodologias científicas. As práticas que têm sido mais freqüentemente objeto de avaliação

são aquelas resultantes da ação social planejada tais como políticas, programas e serviços de saúde.

Entre esses dois extremos opostos do espectro da avaliação estaria uma multiplicidade de possibilidades de avaliação para a gestão ou avaliação administrativa em saúde. A realização de um inquérito populacional para avaliar a cobertura vacinal em um município determinado pode, por exemplo, ser uma avaliação administrativa que estaria próxima a pesquisa avaliativa por recorrer a metodologia científica. Contudo, pode ser considerada como pesquisa avaliativa se vier a preencher uma lacuna no conhecimento sobre a situação do controle de determinado problema de saúde como a poliomielite, plausível de ser atribuída a essa ação (VIEIRA-DA-SILVA e outros, 1997).

Esse espectro da avaliação tem sido reconhecido por alguns autores, tanto em relação ao objeto da avaliação quanto ao método e abordagem. Alguns deles referem-se a avaliação presente no “mundo não acadêmico” e na vida privada, nesta incluindo outros setores das práticas sociais como degustação de vinho e astrologia (PATTON, 1997; SCRIVEN, 2001) ou avaliação “pessoal” (SCRIVEN, 1991).

Quanto ao objeto, alguns autores, apesar de reconhecerem a diversidade do campo da avaliação enfocam principalmente os métodos e técnicas relacionados a avaliação de programas sociais (SCRIVEN, 1991; ROSSI; FREEMAN, 2004; PATTON, 1997). Existe uma preocupação em estabelecer diferenças entre o que seria pesquisa e o que seria avaliação normativa ou administrativa, havendo inclusive variações acerca daquilo que se considera como pesquisa (CONTANDRIOPOULOS; CHAMPAGNE e outros, 1997; SCRIVEN, 2001). Alguns autores tratam exclusivamente da pesquisa avaliativa, embora reconheçam que a avaliação é uma arte que envolve outras habilidades além da metodologia científica (ROSSI; FREEMAN, 2004).

Patton, refere-se a enorme diversidade existente ao interior do campo da avaliação e adverte que as tentativas de reduzir essa complexidade a algumas categorias corre o risco de excessiva simplificação. Entretanto, para fins práticos, esse autor discute três objetivos primários para avaliação de programas: a) para fazer julgamentos; b) para facilitar o seu desenvolvimento c) para contribuir com o conhecimento (PATTON, 1997).

Outra dicotomia, freqüente entre os autores de língua inglesa, aparece como a oposição entre avaliação somativa (*summative*) e formativa (*formative*). A avaliação somativa de um programa é aquela conduzida após o término do

programa e tem por objetivos prestar contas a uma audiência externa (órgão financiador) ou gestor. Esse tipo de avaliação distingue-se da denominada avaliação formativa que é freqüentemente conduzida durante a implementação de um programa e que tem por objetivo apoiar o desenvolvimento do mesmo (SCRIVEN, 1991). Em ambos os casos, a definição do “foco da avaliação” requer o diálogo entre os especialistas em avaliação e os clientes (STECHEER; DAVIS, 1987).

Mesmo no interior de subáreas temáticas e/ou movimentos do campo da avaliação esse espectro é sempre referido. O consenso de Gothenburg sobre avaliação de impactos em saúde considerou três possibilidades metodológicas: a) estimativa rápida de impacto (*rapid health impact appraisal*); b) análise do impacto em saúde (*health impact analysis*) e c) revisão de impactos em saúde (*health impact review*) (WHO, 1999).

A TEORIA E A DIVERSIDADE DAS ABORDAGENS

As primeiras décadas de constituição do campo da avaliação foram marcadas por grandes estudos com ênfase no desenho e no método (SUCHMAN, 1967). Posteriormente, com a limitação do poder explicativo dessas investigações, ao lado do seu custo e complexidade, alternativas metodológicas foram buscadas nas meta-análises, nos estudos qualitativos (COOK; REICHARDT, 1986 (1982); PATTON, 1987) e naqueles orientados pela teoria do programa (CHEN, 1990).

A teoria de que se fala não são as grandes teorias sociais, e sim aquelas referentes ao programa, seu contexto e interações (HOUSE, 2001). Chen define teoria como um conjunto de pressupostos, princípios e ou proposições que explicam ou guiam a ação social. Em outras palavras, a teoria do programa deveria especificar tudo que deve ser feito para alcançar os objetivos desejados, quais os impactos que podem ser antecipados e como isso se produz. Esse autor ainda distingue as teorias *causais* ou *descritivas* que teriam como objetivo explicar as relações entre a intervenção, a implementação e os resultados daquelas *normativas* ou *prescritivas* que estariam voltadas para definir a imagem-objetivo, ou seja, como *devem ser* as intervenções, o processo de implementação do programa e os resultados. A teoria necessária para realizar uma avaliação envolveria, na sua opinião, tanto a dimensão causal quanto aquela normativa (CHEN, 1990).

A ênfase na teoria como estratégia metodológica privilegiada para a avaliação de programas remete à elaboração de modelos lógicos ou teóricos⁴ como componente necessário desse processo (MERCER; GOEL, 1994; CONTANDRIOPOULOS; CHAMPAGNE e outros, 2000).

O que se quer chamar a atenção aqui, em relação à teoria, diz respeito a sua relação com a polissemia conceitual e diversidade de abordagens existentes para a avaliação na bibliografia especializada sobre o tema. É constatação recorrente que cada autor que escreve sobre o tema propõe uma nova matriz (*framework*) ou modifica aquelas preexistentes com variações, ou acrescenta novos atributos dos programas a serem avaliados ou define de forma diferente termos antigos (VIEIRA-DA-SILVA e FORMIGLI, 1994; CONTANDRIOPOULOS; CHAMPAGNE e outros, 1997). Desde 1981 que Scriven edita o “Thesaurus” da avaliação, numa referência ao dicionário equivalente da língua inglesa (SCRIVEN, 1991). As diversas matrizes existentes para a avaliação do “desempenho de sistemas de saúde” (KNOWLES e outros, 1997; WHO, 2000; CANADA, 2001; HURST; HUGHES, 2001) são mais um exemplo dessa diversidade e do pequeno grau de consenso existente no campo da avaliação.

Essa diversidade reflete, por um lado, o estágio incipiente de desenvolvimento da avaliação enquanto disciplina científica e a natureza e abrangência do campo da avaliação que comporta uma multiplicidade de possibilidades de recorte do real. Em outras palavras, tendo em vista que as opções técnicas são também opções teóricas (BOURDIEU, 1989), a multiplicidade de formas de definir as abordagens, dimensões e atributos para a avaliação refletem, em alguma medida, as escolhas teóricas dos autores que são tantas quanto os possíveis “pontos de vista” que correspondem a posições que os autores ocupam no espaço social e campos a que pertencem, sua formação intelectual e trajetória social.

ABORDAGENS, NÍVEIS, DIMENSÕES, CRITÉRIOS E INDICADORES PARA A AVALIAÇÃO

Como não há método separado da prática de pesquisa (BOURDIEU e outros, 1999), também para a avaliação isso é verdadeiro. Cada avaliação é um caso particular que requer criatividade por parte do investigador/avaliador na formulação da melhor estratégia, na seleção da abordagem na

definição de níveis e atributos, bem como na seleção de critérios, indicadores e padrões. Principalmente na hipótese de tratar-se de uma avaliação voltada para a gestão dos serviços, há que incorporar os agentes do programa (*stake-holders*) em muitas das definições do foco da avaliação (GUBA; LINCOLN, 1989; PATTON, 1997; FETTERMAN, 1997). A sistematização que se segue pretende auxiliar o avaliador a tomar conhecimento do leque de possibilidades para facilitar o processo de delimitação do objeto e de formulação do projeto de avaliação.

ABORDAGENS

Numa primeira aproximação com o objeto a ser avaliado, poderíamos recorrer à classificação das abordagens possíveis feita por (DONABEDIAN, 1988), que é utilizada e revisitada quase quarenta anos depois (HANDLER e outros, 2001). Com base no enfoque sistêmico e principalmente preocupado em avaliar a qualidade do cuidado médico, Donabedian sistematizou diversas propostas de abordagem na tríade: estrutura-processo-resultados. A estrutura diria respeito aos recursos materiais, humanos e organizacionais, o processo a tudo aquilo que medeia a relação profissional-usuário e os resultados estariam relacionados tanto com o produto das ações (consultas, exames, visitas) quanto com a modificação no estado de saúde de indivíduos e da população. Embora ela tenha utilidade prática na primeira aproximação com o objeto da avaliação, a abrangência dessa classificação e a redução que ela opera do real têm sido apontadas como suas principais limitações (VIEIRA-DA-SILVA; FORMIGLI, 1994).

NÍVEIS E OBJETO DA AVALIAÇÃO

Para prosseguir na delimitação do objeto a ser avaliado é necessário definir quais os níveis da realidade que se quer estudar. O campo da saúde é formado por instituições públicas e privadas que desenvolvem ações de promoção, prevenção e cura voltadas para a população e que podem ser abordadas, entre outras possibilidades, a partir dos seguintes níveis (Figura 2)⁵ :

- a) Ações – diz respeito às ações de promoção, prevenção e cura desenvolvidas pelos agentes individualmente;

- b) Serviços – corresponde a um grau de maior complexidade de organização das ações onde diversos agentes se articulam para desenvolver atividades, coordenadas ou não, voltadas para um grupo etário ou problema de saúde. Como exemplo pode-se citar um serviço para o atendimento pediátrico ou um serviço de saúde ocupacional;
- c) Estabelecimentos – correspondem a unidades sanitárias de diferentes níveis de complexidade tais como centros de saúde, hospitais, policlínicas.
- c) Sistema⁶ – seria o nível mais complexo de organização das práticas que envolveria todos os outros e sua coordenação. Pode corresponder a um sistema municipal, estadual ou nacional.

A avaliação de políticas, programas e projetos pode recortar todos os níveis do sistema de saúde. As políticas de saúde podem ser consideradas como a ação ou omissão do Estado enquanto resposta social, diante dos problemas de saúde e seus determinantes, bem como da produção, distribuição e regulação de bens e serviços e ambientes que afetam a saúde dos indivíduos e da coletividade (PAIM, 2002; PAIM, 2003⁷). Dessa forma, a avaliação de políticas pode incluir, tanto análises sobre a natureza do Estado e do poder político envolvidos na sua formulação quanto estudos sobre os programas⁸ relacionados com a sua operacionalização. Por fim, embora a avaliação de políticas frequentemente envolva a avaliação de programas, a distinção entre essas duas dimensões das práticas pode ser necessária para fins analíticos (VIEIRA-DA-SILVA, 1999)⁹.

ATRIBUTOS OU CARACTERÍSTICAS

A seleção das características ou atributos das práticas, serviços, programas, estabelecimentos ou do sistema de saúde a serem avaliadas contribui para o aprofundamento no processo de definição do foco da avaliação. Aqui é onde existe a maior diversidade terminológica e polissemia conceitual. As opções ora apresentadas representam uma atualização de proposição anterior (VIEIRA-DA-SILVA; FORMIGLI, 1994) onde buscou-se incorporar conceitos com certo grau de consenso na literatura revisada que guardassem coerência entre si e com um modelo teórico (Figura 3).

Em resposta ou não a necessidades de saúde, os indivíduos desenvolvem necessidades de serviços de saúde (PAIM, 1982). Essas necessidades podem ser objeto de graus distintos de controle, a saber, controle de causas, de riscos ou de danos que dependem do estágio de desenvolvimento da história de um agravo à saúde e das características do modelo assistencial (PAIM, 1999). Para cada um desses níveis de controle existem diversas possibilidades de intervenção (Quadro 1) cujas características ou atributos têm sido consideradas como relevantes para a avaliação, podendo ser agrupados de acordo com as seguintes características, entre outras possibilidades (Figura 3):

- a) Relacionados com a disponibilidade e distribuição social dos recursos:
 - Cobertura,
 - Acessibilidade,
 - Eqüidade.
- b) Relacionados com o efeito das ações:
 - Eficácia,
 - Efetividade,
 - Impacto.
- c) Relacionados com os custos e produtividade das ações:
 - Eficiência.
- d) Relacionados com a adequação das ações ao conhecimento técnico e científico vigente:
 - Qualidade técnico-científica.
- e) Relacionados com a adequação das ações aos objetivos e problemas de saúde:
 - *Direcionalidade e consistência* (análise estratégica).
- f) Relacionados com o processo de implantação das ações:
 - avaliação do grau de implantação e (ou) avaliação de processo,
 - análise de implantação – estudos que investigam as relações entre o grau de implantação, o contexto e os efeitos das ações.

g) Características relacionais entre os agentes das ações:

- usuário x profissional (percepção dos usuários sobre as práticas, satisfação dos usuários, aceitabilidade, acolhimento, respeito a privacidade e outros direitos cidadãos),
- profissional x profissional (relações de trabalho e no trabalho),
- gestor x profissional (relações sindicais e de gestão).

A primeira preocupação do gestor em relação a uma intervenção sanitária está relacionada com a magnitude da oferta das ações e serviços. Em que medida os recursos existentes são suficientes para atender às necessidades da população-alvo? Essa pergunta pode ser respondida através da avaliação de *cobertura* que é aquela que mede a proporção da população que se beneficia do programa ou intervenção (ROSSI; FREEMAN, 2004). A cobertura pode ser pensada em relação a intervenções de diversos tipos, desde aquelas voltadas para a promoção da saúde ou para o controle de causas até aquelas vinculadas ao cuidado propriamente dito, ou seja, ao controle de danos (Quadro 1). A cobertura pode ser considerada como potencial quando o que se mede são os recursos materiais ou humanos disponíveis para a realização de determinada ação. Nesse caso, ela é superponível ao conceito de “oferta de serviços”. Já a cobertura real ou utilização de serviços corresponde a proporção da população que efetivamente se beneficiou de determinada intervenção, que de fato *usou* os serviços (SOBERÓN, 1988). A cobertura pode ser calculada tendo por referência um sistema estruturado para o atendimento à demanda espontânea ou ser pensada em relação à oferta organizada¹⁰ de ações programáticas (Figura 3).

A *acessibilidade*, embora seja um conceito relacionado com o de cobertura, não é equivalente. Ela tem sido definida como uma relação entre os recursos de poder dos usuários e os obstáculos colocados pelos serviços de saúde (FRENK, 1992). Os recursos de poder podem ser de natureza econômica, social e cultural. Já os obstáculos podem ser geográficos (distância, transporte), organizacionais (existência de filas, tempo de espera injustificável, natureza do acolhimento) e econômicos.

O *acesso* pode ser confundido com a *utilização de serviços* quando só se quer analisar a obtenção do cuidado pelo indivíduo que dele necessita. Quando existem dificuldades importantes para essa *utilização*, o conceito de *acessibilidade* relacionado aos obstáculos existentes à obtenção dos servi-

ços, por parte dos usuários, passa a ter grande valor, como é o caso do Brasil. Segundo a Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada em 1998, 5 milhões de pessoas referiram ter necessitado, mas não procuraram um serviço de saúde, alegando principalmente falta de recursos financeiros (obstáculo econômico). Estimativas desse mesmo estudo indicam que 755.521 pessoas que procuraram atendimento não foram atendidas por motivos diversos: ou não conseguiram vaga ou senha (obstáculos organizacionais ou de oferta), ou não havia serviço especializado (oferta), ou não havia médico atendendo ou esperaram muito e desistiram (obstáculos organizacionais). Além disso, o acesso a consultas médicas aumenta expressivamente com a renda (recursos de poder econômicos) (BRASIL-IBGE, 2000). Esses dados revelam o sofrimento dos usuários muitas vezes obrigados a aguardar em longas filas nas portas dos serviços de saúde. Por outro lado, o tempo de espera para o acesso a procedimentos especializados pode interferir com o próprio efeito do cuidado se o mesmo não é feito tempestivamente.

Já o conceito de *equidade*¹¹ aqui será usado no sentido que lhe é dado por Perelman, ou seja, como um instrumento da justiça para resolver as contradições entre as diversas fórmulas da justiça formal ou abstrata (PERELMAN, 1996). O que equivale a dizer que devemos tratar desigualmente desiguais e priorizar para a intervenção sanitária grupos sociais com maiores necessidades de saúde. A equidade recorta e relaciona-se com quase todos os demais atributos. Pode-se pensar em equidade no acesso, no acolhimento, na qualidade e na efetividade. A oferta organizada de ações voltadas para enfrentar determinado problema de saúde pode ampliar a acessibilidade da população aos diversos níveis do cuidado e, dessa forma, ampliar a cobertura real das referidas ações.

A segunda e importante característica das práticas de saúde diz respeito aos possíveis efeitos que as mesmas possam ter sobre o estado de saúde dos indivíduos, grupos ou populações. A distinção entre *eficácia* como efeito de uma intervenção em situação experimental e *efetividade* como o efeito da mesma em sistemas operacionais vem se mantendo ao longo dos anos (DONABEDIAN, 1990; MAYNARD; MCDAID, 2003). Já o conceito de impacto tem sido usado frequentemente com o mesmo sentido dado a efetividade, ou seja, relacionado ao efeito de intervenções em sistemas reais e não em situações experimentais da pesquisa. Contudo, o conceito de impacto tem sido também usado para designar o efeito de uma intervenção

em relação a grandes grupos populacionais ou em grandes intervalos de tempo. Exemplificando, seria necessário que determinada tecnologia fosse *eficaz* em ensaios experimentais para que, ao ser utilizada em sistemas operacionais possa ser *efetiva*, e, se associada a uma cobertura elevada, durante determinado período, tenha a possibilidade de causar *impacto* sobre o nível de saúde de uma população com redução de morbidade. Uma vacina que obtém elevada eficácia em ensaios experimentais pode ter baixa efetividade ao ser administrada em usuários de um Centro de Saúde devido a problemas da rede de frio, por exemplo. Já para o controle de viroses para as quais exista uma vacina *eficaz*, é necessário que os serviços atinjam elevadas coberturas vacinais em todo o país, ou seja, sejam *efetivos*, para então se falar em *impacto* da medida.

A eficiência tanto tem sido usada como medida da produtividade do sistema como da sua relação com os custos (HARTZ; POUVOURVILLE, 1998). A incorporação de avaliações econômicas na definição das prioridades relacionadas com a saúde das populações tem sido feita para fazer face aos crescentes custos da atenção médico-hospitalar e também para garantir a equidade de ações de promoção à saúde e reformas na organização dos sistemas de saúde (MCDAID, 2003).

A qualidade técnico-científica do cuidado no sentido que lhe foi dado por Vuori (1982), corresponde à adequação das ações ao conhecimento técnico e científico vigente. Ou seja, em que medida as tecnologias que estão sendo usadas são aquelas consideradas como mais indicadas e eficazes pela comunidade científica?

No caso das ações conduzidas a partir de algum grau de planejamento a *análise estratégica*¹² (CAMPOAMOR, 1990; CONTANDRIOPOULOS; CHAMPAGNE e outros, 1997) auxilia a responder a pergunta acerca da consistência e direcionalidade do plano (MATUS, 1993). Ou seja, em que medida as atividades propostas são pertinentes, tendo em vista os objetivos do programa e os problemas diagnosticados?

Já as *análises de implantação* são muito relevantes para a avaliação de programas de saúde porque permitem ao investigador estudar os processos relacionados à operacionalização dos mesmos. As avaliações de efetividade, conduzidas, muitas vezes a partir de estratégias do tipo “caixa preta¹³”, chegam a resultados bastante diferenciados com problemas de generalização de resultados (DENIS; CHAMPAGNE, 1997; HOUSE, 2001). Abrir a “caixa pre-

ta” não apenas tem elucidado problemas na operacionalização dos programas como tem sido útil para a formulação de recomendações voltadas para o aperfeiçoamento dos mesmos (HULSCHER e outros, 2003). Os processos, a maneira pela qual os programas são implementados, podem ser tão importantes quanto os resultados. Em algumas circunstâncias, o processo de implantação de programas de reorganização da atenção à saúde pode envolver a participação dos diversos profissionais, promover a integração entre os mesmos e os gestores de tal forma, que o próprio desenvolvimento das relações interpessoais pode passar a ser um resultado relevante (PATTON, 2002)¹⁴.

Por último, o bloco dos atributos denominados como “relacionais”, que podem ou não fazer parte das análises de implantação, cria toda uma nova perspectiva para a avaliação em saúde. Corresponde a valorizar as características do acolhimento, a garantia dos direitos individuais à privacidade e as características das relações no trabalho como componentes da qualidade da atenção e promoção da saúde.

As intervenções de saúde variam conforme o modelo assistencial ou modelos assistenciais dominantes em diferentes contextos. As características acima relacionadas podem ser pensadas tanto para a atenção à demanda espontânea quanto para a oferta organizada voltada para o controle de riscos e causas (PAIM, 1999).

ESTRATÉGIAS E DESENHOS

As estratégias de pesquisa para a avaliação correspondem às possíveis maneiras de mobilizar recursos teóricos e técnicos visando responder à pergunta de investigação de forma a evitar vieses e também responder a hipóteses rivais. O conceito de desenho será usado para designar aquelas estratégias mais estruturadas, relacionadas principalmente com a investigação em epidemiologia (ALMEIDA-FILHO e ROUQUAYROL, 1999).

Os desenhos possíveis para a pesquisa avaliativa têm sido propostos a partir de uma comparação com o desenho experimental ou de intervenção, tido como tipo ideal, capaz de responder aos principais desafios à validade interna (CAMPBELL; STANLEY, 1966). Esses autores discutem 12 fatores que ameaçam a validade da pesquisa em cerca de 16 diferentes desenhos. Teoricamente, o desenho experimental seria capaz de dar conta da maioria desses possíveis vieses tais como: história, maturação, familiaridade ao teste,

instrumentação, regressão à média, seleção, mortalidade experimental entre outros (CAMPBELL; STANLEY, 1966). Os desenhos experimentais correspondem aos “ensaios clínicos” ou estudos de intervenção na epidemiologia (Quadro 2).

A crítica às limitações do desenho experimental para avaliação de programas dizem respeito a problemas éticos e operacionais (SMITH, 1980). A escolha aleatória de participantes para um estudo geralmente encontra resistências com recusas frequentes tanto devido ao caráter arbitrário dessa designação quanto devido as diferenças na natureza e atratividade das intervenções. Por outro lado, os estudos experimentais são caros e demandam longo tempo. Do ponto de vista ético, não se admite observar um grupo de pessoas com um problema sem aportar nenhuma solução para o mesmo (SMITH, 1980). Particularmente, no que diz respeito às avaliações em saúde, os desenhos observacionais são preferidos em detrimento dos experimentais, tendo em vista principalmente questões éticas, bem como a complexidade das intervenções de saúde coletiva e sua relação com o contexto (ALMEIDA-FILHO; ROUQUAYROL, 1999).

Como uma alternativa às limitações da ênfase no método e nos desenhos do tipo caixa preta, experimentais, têm sido propostas estratégias complementares buscando explicar variações nos resultados a partir das variações na implantação e no contexto (DENIS; CHAMPAGNE, 1997; HARTZ, 1997; CONTANDRIOPOULOS e outros, 2003; FRIAS; HARTZ, 2003; VIEIRA-DA-SILVA e outros, 2003), com ênfase na teoria (CHEN, 1990) e com recurso a técnicas não estruturadas ou qualitativas (COOK; REICHARDT, 1986; PATTON, 1987; GUBA; LINCOLN, 1989; FETTERMAN, 1997).

O debate quantitativo x qualitativo é considerado como superado por alguns autores, tendo em vista que limites e possibilidades são hoje reconhecidos tanto em relação ao uso de técnicas quantitativas quanto para aquelas qualitativas (COOK, 1997; HOUSE, 2001). Alguns, porém, ainda consideram como atual o debate argumentando que se tratam de divergências filosóficas (SMITH, 2001). Em outras disciplinas do campo científico, essa é considerada como uma falsa questão (BOURDIEU, 1989). O que é considerado, por esse último autor, como relevante para o processo de investigação, é a construção do objeto e a mobilização de todas as técnicas possíveis para analisá-lo (BOURDIEU, 1998).

No caso particular da avaliação, existem muitos objetos pré-construídos dado seu caráter predominantemente instrumental, principalmente nas

avaliações formativas e voltadas para a gestão. Contudo, na pesquisa avaliativa, há necessidade e existe um grande espaço de possibilidades de construção do objeto, principalmente nas análises de implantação.

Menos com a intenção de apresentar um “livro de receitas” e mais com o objetivo de sistematizar possibilidades, para orientar escolhas práticas, pode-se classificar as estratégias possíveis em estruturadas ou “pronto para usar” (*prêt-à-porter*)¹⁵ e semi estruturadas¹⁶ ou “sob medida” (*sur mesure*) (HARTZ, 1999). Entre as estratégias estruturadas podem ser incluídos os clássicos desenhos experimentais, os quase-experimentais e mesmo alguns considerados como não experimentais (CAMPBELL; STANLEY, 1966). Os desenhos epidemiológicos que têm sido usados para avaliação em saúde podem ser considerados como estruturados e apresentam uma equivalência com aqueles propostos por Campbell e Stanley (Quadro 2).

Já os pouco estruturados geralmente são estudos de caso¹⁷ que, a depender dos níveis de análise e da profundidade, podem mobilizar diversas técnicas de coleta de dados qualitativos como: observação etnográfica, grupos focais, entrevistas em profundidade, histórias de vida e análise documental (COOK; REICHARDT, 1986 (1982); YIN, 1994) (PATTON, 1987; GUBA; LINCOLN, 1989).

A escolha da estratégia relaciona-se com o processo anterior de delimitação do foco e de construção do objeto da avaliação. Avaliar a dimensão relacional das práticas requer a realização de estudos de caso e o uso de técnicas qualitativas (HADAD e outros, 2002). Da mesma forma nas análises de implantação a resposta à pergunta acerca das relações entre o grau de implantação e os contextos políticos e organizacionais requer a realização de análises históricas, entrevistas em profundidade e a elaboração de modelos teóricos causais (DENIS; CHAMPAGNE, 1997; HARTZ e outros, 1997, 2003; PATTON, 2002). Esses dois tipos de avaliação são muito úteis para o gestor no que diz respeito a geração de informações voltadas para as retificações necessárias nos programas em curso (HULSCHER; LAURANT e outros, 2003).

A avaliação de cobertura e de utilização de serviços impõe a quantificação e o uso de desenhos estruturados como o inquérito (VIEIRA-DASILVA e outros, 1995; TRAVASSOS e outros, 2002). Para as avaliações de eficácia de tecnologias os desenhos indicados ainda são os experimentais ou quase-experimentais observacionais do tipo estudos ecológicos, caso-controle e coortes (BARRETO e outros, 2002).

Já em relação a avaliação da efetividade de programas há uma tendência crescente ao uso de síntese de pesquisas como uma estratégia prioritária para responder a perguntas sobre o efeito de tecnologias ou programas em sistemas operacionais (COOK, 1997). Por outro lado, diante da complexidade da avaliação da efetividade de programas em contextos diferenciados, propõe-se o uso articulado entre diversas investigações com desenhos tradicionais ou não, tendo em vista a limitação dos desenhos quando considerados de forma isolada (PAWSON; TILLEY, 2001). A ampliação da base de evidências assim conseguida, ao lado da utilização da teoria e de desenhos sintéticos pode permitir a realização de inferências causais com maior segurança (CHEN, 1990). Por fim os estudos considerados como de desenvolvimento de novas tecnologias para intervenção ou avaliação requerem uma ênfase no pólo teórico e síntese de conhecimento (CONTANDRIOPOULOS e outros, 1994).

CONSIDERAÇÕES FINAIS

O campo da avaliação expandiu-se consideravelmente no final do século XX tanto em produção científica quanto no que diz respeito a sua institucionalização. As suas principais características são a diversidade conceitual e terminológica, a pluralidade metodológica e a multiplicidade de questões consideradas como pertinentes. Persiste uma importante lacuna relacionada com a incorporação do conhecimento produzido no plano da investigação pelos profissionais de saúde e pelos gestores.

Há um crescente consenso que as diferenças nas abordagens e estratégias metodológicas para a avaliação dependem dos objetos e circunstâncias (ROGERS, 2001). O julgamento adequado sobre o valor de uma intervenção requer muitas vezes a mobilização de diferentes estratégias metodológicas, visita ao conhecimento anteriormente produzido e desenvolvimento teórico-conceitual (PAWSON; TILLEY, 2001). Em síntese, a avaliação pode produzir informação tanto para a melhoria das intervenções em saúde como para o julgamento acerca da sua cobertura, acesso, equidade, qualidade técnica, efetividade, eficiência e percepção dos usuários a seu respeito. Com essa finalidade pode e deve mobilizar as estratégias e técnicas necessárias da epidemiologia e das ciências humanas, sem preconceitos e sem interdições.

Quadro 1

Possíveis relações entre abordagens, características selecionadas das práticas e níveis de controle

<p>Abordagens</p> <p>Características das práticas</p>	<p>Estruturas</p>	<p>Processos</p>	<p>Resultados (Efetividade /impacto)</p> <p>Níveis de controle</p> <p>Controle de causas</p>
<p>Cobertura</p>	<p>Cobertura potencial de uma ação educativa relacionada com a promoção da saúde</p> <p>(ex: proporção de panfletos disponíveis em relação a população-alvo relacionados com os benefícios de fazer exercícios físicos)</p>	<p>Cobertura real dessa mesma ação educativa relacionada com a promoção da saúde</p> <p>(ex: proporção de panfletos efetivamente distribuídos em relação à população-alvo)</p>	<p>Promoção da saúde</p>
<p>Cobertura</p>	<p>Cobertura potencial de ações relacionadas ao controle da qualidade da água</p>	<p>Cobertura real dessas mesmas ações. Organização adequada do acolhimento em Centros de Saúde</p>	<p>Controle de riscos</p> <p>Eliminação de agentes potencialmente causadores de doenças de transmissão hídrica</p>
<p>Acessibilidade</p>	<p>Distribuição geográfica adequada dos Centros de Saúde. Equipamentos necessários à realização de exames complementares</p>	<p>Cuidado adequado ao diagnóstico</p>	<p>Controle de danos</p> <p>Pessoas com agravos à saúde que tenham tido acesso ao cuidado adequado podem curar-se do agravo</p>
<p>Qualidade técnica</p>			<p>Controle de danos</p> <p>Cura</p>

Quadro 2

Tipo de Estratégia/Desenho

	Experimental	Quase Experimental	Não Experimental
Estruturados	Pré-Teste Pós-teste Com controle 01 X 02 03 X 04 Campbell(1963)	Séries Temporais ...0 ₁ 0 ₂ 0 ₃ x 0 ₄ 0 ₅ 0 ₆ Campbell(1963)	Um grupo Pós-teste X 01 Pré-Teste Pós-teste 01 X 02 Campbell(1963)
	Intervenção Lilienfeld (1976)	Séries Temporais Caso-Controlle Coortes Lilienfeld (1976)	Estudos transversais Lilienfeld (1976)
	Intervenção Almeida Filho; Rouquayrol (1999)	Observacional Agregados Individual Longitudinal Almeida-Filho; Rouquayrol (1999)	Estudos transversais Almeida-Filho; Rouquayrol (1999)
Semi-estruturados	Estudos de caso com análise histórica -	Estudo comparado de casos múltiplos Estudos centrados na teoria e desenhos sintéticos (CHEN, 1990)	Estudos de caso (YIN,1994) Pesquisa de desenvolvimento (CONTANDRIOPOULOS, 1994)

Figura 1

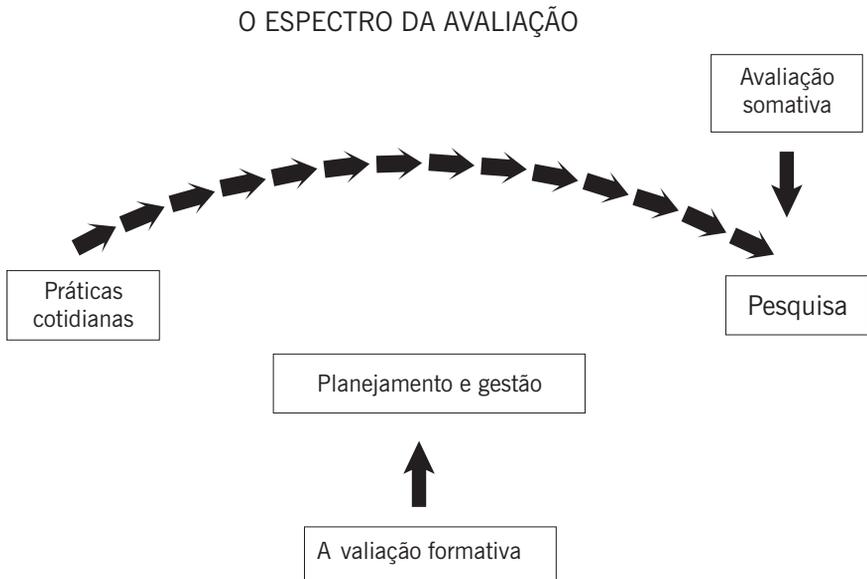


Figura 2

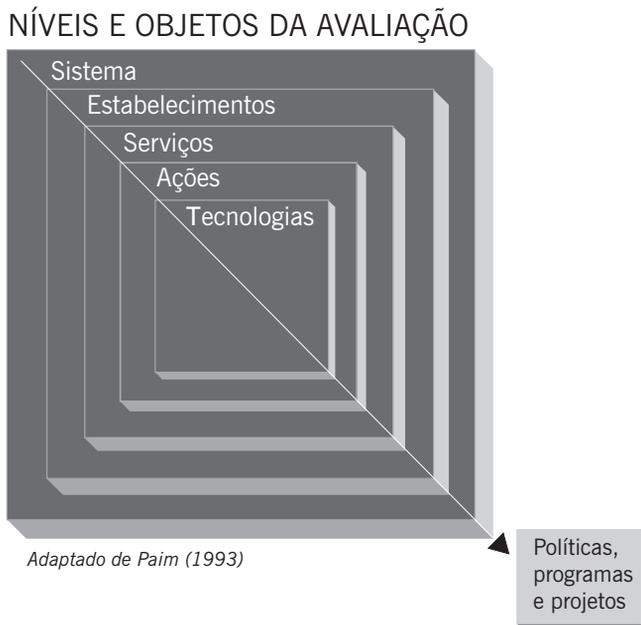
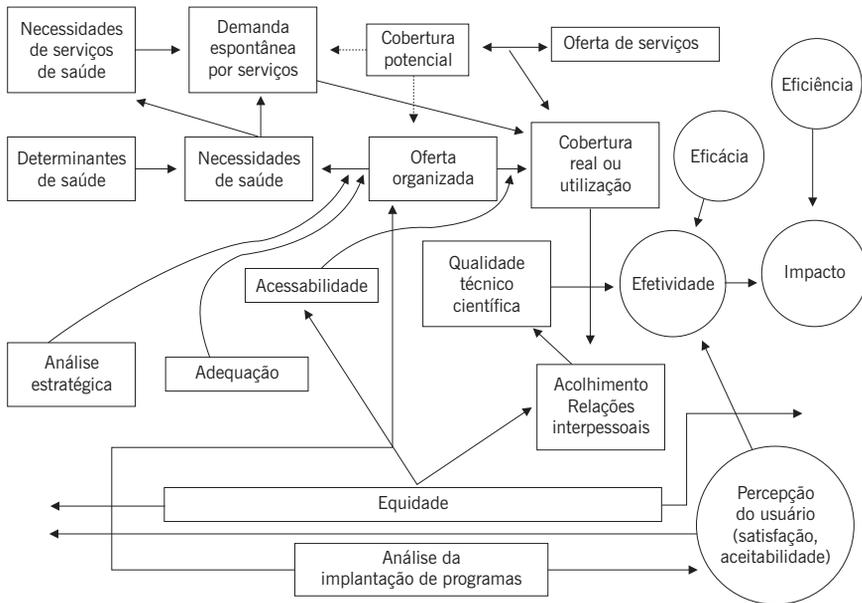


Figura 3
Atributos ou Características para Avaliação em Saúde



NOTAS

¹ Para Contandriopoulos, uma intervenção é um conjunto de meios (físicos, humanos, financeiros e simbólicos) organizados em um contexto específico, em um dado momento, para produzir bens ou serviços para modificar uma situação problemática (CONTANDRIOPOULOS, 1997).

² O conceito de práticas sociais usado aqui é aquele desenvolvido por Bourdieu, segundo o qual, as práticas sociais são a relação entre uma situação e um *habitus*. Para melhor compreensão dos conceitos de campo e *habitus* sugere-se a leitura dos livros *La Distinction* (1979) e *Le sens pratique* (1980) deste autor.

³ Embora o termo “avaliativo” seja um anglicismo o seu uso vem sendo progressivamente incorporado aos textos acadêmicos sobre a avaliação. A expressão que consta do *Dicionário Aurélio*, “avaliatório”, é menos utilizada embora também correta.

⁴ Uma análise mais detalhada sobre modelos lógicos e teóricos é objeto do capítulo 2 do presente livro.

⁵ Tipologia adaptada de (PAIM, 1993).

⁶ Ver discussão detalhada sobre a avaliação de sistemas em (VIEIRA-DA-SILVA, 1999).

⁷ Paim, J. S. Políticas de saúde no Brasil. *Epidemiologia e Saúde*. Rouquayrol; Almeida-Filho. Rio de Janeiro: MEDSI, 2003 (no prelo).

⁸ Os programas correspondem a um grau de organização das práticas onde há definição de objetivos, atividades e recursos para resolução de problemas definidos de saúde. O programa pode ou não estar sendo desenvolvido ao interior de um serviço, estabelecimento ou sistema de saúde.

⁹ Diversas outras possibilidades de delimitação de níveis existem. Nos capítulos 4, 5 e 6 do presente livro o leitor verá diferentes níveis selecionados para a realização de investigações avaliativas.

¹⁰ Os conceitos de demanda espontânea e oferta organizada são aqueles utilizados por Paim (1999).

¹¹ Uma análise detalhada do significado do conceito de equidade é feita no ensaio “Vieira-da-Silva; Almeida-Filho, 2000. Distinção, diferença, desigualdade, iniquidade e a saúde: uma análise semântica. VI Congresso da ABRASCO. Salvador, Bahia.”

¹² A “análise estratégica” para Contandriopoulos e Campoamar tem um sentido distinto daquele que é usado pela OPS no documento “*Formulacion de Políticas de Salud*” (CPPS/OPS, 1975). Neste último documento, a definição de estratégias relaciona-se com a análise da viabilidade política para a implementação do plano enquanto que a análise das relações entre objetivos e proposições é denominada análise de coerência (CPPS/OPS, 1975).

¹³ A expressão “caixa preta” tem sido usada por diversos autores numa referência aos desenhos que só investigam os resultados das intervenções não se preocupando com os processos e contextos.

¹⁴ Diversos exemplos de análise de implantação de programas são descritos no presente livro. No capítulo 7, Contandriopoulos e colaboradores descrevem uma experiência canadense de integração de serviços de saúde que foi objeto de uma investigação sobre a sua implantação. Já os capítulos 5, 4 e 6 referem-se a investigações realizadas no Brasil sobre a implantação de programas e políticas de abrangência diversa tais como: a assistência farmacêutica (FRIAS; HARTZ), a intersectorialidade (COELHO; VIEIRA-DASILVA) e a descentralização da gestão do SUS (VIEIRA-DASILVA; HARTZ e outros).

¹⁵ A analogia entre os desenhos estruturados da pesquisa e a costura do tipo “pronta entrega” (*pret-à-porter*) por um lado e os estudos semi-estruturados e a costura feita “sob medida” (*sur mesure*) por outro foi feita por Hartz em debate sobre a institucionalização da avaliação (HARTZ, 1999).

¹⁶ A expressão “semi-estruturadas” aqui não tem o sentido de pouco rigor e sim refere-se à flexibilidade possível nessas estratégias de investigação.

¹⁷ Os estudos de caso podem ser únicos ou múltiplos, com uma ou diversas unidades ou níveis de análise (YIN, 1994).

REFERÊNCIAS

ALMEIDA-FILHO, N.; ROUQUAYROL, M.Z. Desenhos de Pesquisa em Epidemiologia. *Epidemiologia e Saúde*. Almeida-Filho, N. e Rouquayrol, M.Z. São Paulo; Belo Horizonte, Editora Médica e Científica Ltda. MEDSI. 1999.

BARRETO, M.; RODRIGUES, L.; CUNHA, S.; PEREIRA, S.; HIJJAR, M.; ICHIHARA, M.; DE BRITO, S.C.; I.,D. "Design of the Brazilian BCG-REVAC trial against tuberculosis: a large, simple randomized community trial to evaluate the impact on tuberculosis of BCG revaccination at school age." *Control Clin Trials* v. 23, n. 5, p. 540-553, 2002.

BOURDIEU, P. *La distinction. Critique social du jugement*. Paris: Minuit, 1979.

BOURDIEU, P. *O poder simbólico*. Rio de Janeiro, Lisboa: DIFEL, Bertrand Brasil, 1989.

BOURDIEU, P.; CHAMBOREDON, J.C.; PASSERON, J.C. *A profissão do sociólogo: preliminares epistemológicas*. Petrópolis: Vozes, 1999.

BOURDIEU, P.; WACQUANT, L.J.D. *Réponses. Pour une anthropologie réflexive*. Paris: Seuil, 1992.

BRASIL-IBGE. *Pesquisa Nacional por amostra de domicílios. Acesso e utilização de serviços de saúde*:1998. Rio de Janeiro: IBGE: 96. 2000.

CAMPBELL, D.T.; STANLEY, J.C. *Experimental and quasi-experimental designs for research*. Chicago: Rand Mc Nally College Publishing Company (Reprinted from Handbook of Research on Teaching, American Educational Research Association, 1963), 1966.

CAMPOAMOR, N.R. Evaluación de los sistemas locales de salud. *Los sistemas locales de salud*. J. M. Paganini and R. C. Mir. Washington, D.C.: Organizacion Panamericana de La Salud, 1990. p. 463-467.

CANADA. *Health Care in Canada*. Ottawa: Canadian Institute for Health Information, 2001.

CHELIMSKY, E. The coming transformations in evaluation. *Evaluation for the 21st century: a handbook*. C.E.S.W. (eds.). Thousand Oaks-London-New Delhi: Sage Publications, 1997. p. 1-26.

CHELIMSKY, E.; SHADISH, W. R. Preface. *Evaluation for the 21st century: a handbook*. S. Publications. Thousand Oaks-London-New Delhi: Sage Publications: xi-xiii, 1997.

CHEN, H. *Theory – driven evaluations*. Newbury Park: Sage Publications, 1990.

CONTANDRIOPOULOS, A.; CHAMPAGNE, F.; DENIS, J.; AVARGUES, M. "L'évaluation dans le domaine de la santé: concepts et méthodes." *Revue Epidemiologie et Santé Publique*, n. 48, p. 517-539, 2000.

CONTANDRIOPOULOS, A.; CHAMPAGNE, F.; DENIS, J.; PINEAULT, R. Avaliação na área de saúde: conceitos e métodos. *Avaliação em saúde: dos modelos conceituais à prática na análise da implantação de programas*. Rio de Janeiro: FIOCRUZ/Hartz, ZA, 1997. p. 29-48.

CONTANDRIOPOULOS, A.P.; CHAMPAGNE, F.; POTVIN, L.; DENIS, J.L.; BOYLE, P. *Saber preparar uma pesquisa*. São Paulo/Rio de Janeiro: Hucitec/Abrasco, 1994. p. 35-55 / 79-86.

COOK, T. Lessons learned in evaluation over the past 25 years. *Evaluation for the 21st century: a handbook*. C.E.S.W. (eds.). Thousand Oaks: Sage Publications, 1997. p. 30-52.

COOK, T.D.; Reichardt, C.S. *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid: Ediciones Morata. (1986; 1982)

DENIS, J.; CHAMPAGNE, F. Análise da implantação de programas. *Avaliação em saúde: dos modelos conceituais à prática na análise da implantação de programas*. Rio de Janeiro: FIOCRUZ/Hartz, ZA. (ed.). 1997. p. 49-88.

DONABEDIAN, A. The definition of quality: a conceptual exploration. *Explorations in quality assessment and monitoring. v. 1: the definition of quality and approaches to its assessment*. A. Donabedian. Ann Arbor, Michigan: Health Administration Press, 1988. p. 3-31.

DONABEDIAN, A. "The seven pillars of quality." *Arch. Pathol Lab Med*, 114, p. 1115-1118, 1990.

FETTERMAN, D.M. *Empowerment evaluation and accreditation in higher education*. Thousands Oaks, London, New Delhi: Sage Publications, 1997.

FRENK, J.M. El concepto y la medición de accesibilidad. *Investigaciones sobre servicios de salud: una antología*. K. L. OPAS (White, editor principal). Washington, D.C.:OPS, 1992. p. 929-943.

GRÉMY, F. "Evaluation of medical action and of action in public health." *Effective Health Care*, v. 1, n. 1, p. 11-23, 1983.

GUBA, E.; Lincoln, Y. Fourth generation evaluation. *Chapter 1: the coming of age of evaluation*. Newbury Park: Sage Publications, 1989. p. 21-49.

HABERMAS, J. *Teoría de la acción comunicativa I*. Madri: Taurus, 1987.

HADAD, S.; FRANCA, E.; UCHOA, E. "Preventable infant mortality and quality of health care: maternal perception of the child's illness and treatment." *Cad Saude Publica*, v. 18, n. 6, p. 1519-1527, 2002.

HANDLER, A.; ISSEL, M.; TURNOCK, B. "A conceptual framework to measure performance of the public health system." *American Journal of Public Health*, v. 91, n. 8, p. 1235-1239, 2001.

HARTZ, Z.A. "Institucionalizing the evaluation of health programs and policies in France: cuisine internationales over fast food and sur mesure over ready-made." *Cadernos de Saúde Pública*, 1999, p. 229-260.

HARTZ, Z.A.; Champagne, F.; CONTANDRIOPOULOS, A.P.; C, L.M.d. Avaliação do Programa Materno-infantil : análise de implantação em sistemas locais de saúde no Nordeste do Brasil. *Avaliação em saúde: dos modelos conceituais à prática na análise da implantação de programas*. H.Z. (ed.). Rio de Janeiro: FIOCRUZ, 1997.

HARTZ, Z.A.; POUVOURVILLE, G. "Avaliação dos programas de saúde: a eficiência em questão." *Ciência e Saúde Coletiva*, v. 3, n. 1, p. 68-82, 1998.

HOUSE, E.R. "Unfinished Business: causes and values." *American Journal of Evaluation*, v. 22, n. 3, p. 309-315, 2001.

HULSCHER, M.E.J.L.; LAURANT, M.G.H.; GRAL, R.P.T.M. "Process evaluation on quality improvement interventions." *Qual Saf Health Care*, n. 12, p. 40-46, 2003.

- HURST, J.; HUGHES, M.J. Performance Measurement and Performance Management in OECD Health Systems, OECD, 2001.
- KNOWLES, J.C.; LEIGHTON, C.; STINSON, W. Measuring results of Health Sector Reforma for System Performance: a Handbook of Indicators. Bethesda, Maryland, ABT Associates Inc. Development Associates Inc. Harvard School of Public Health, 1997.
- MATUS, C. *Política, Planejamento e Governo*. Brasília: IPEA, 1993.
- MAYNARD, A.; MCDAID, D. "Evaluating health interventions: exploiting the potential." *Health Policy*, n. 63, p. 215-226, 2003.
- MCDAID, D. "Evaluating health interventions in the 21st century: old and new challenges." *Health Policy*, n. 63, p. 117-120, 2003.
- MERCER, S.L.; GOEL, V. "Program evaluation in the absence of goals: a comprehensive approach to the evaluation of a population-based breast cancer screening program." *The Canadian Journal of Program Evaluation*, v. 9, n. 1, p. 97-112, 1994.
- ORGANIZACIÓN PANAMERICANA DE LA SALUD/Centro Panamericano de Planificación de la Salud. *Formulación de Políticas de Salud*. Santiago, 1975. 77 p.
- PAIM, J.S. As ambiguidades da noção de necessidades de saúde. *Introdução à crítica ao planejamento de saúde*. Salvador: Universidade Federal da Bahia-Departamento de Medicina Preventiva, 1982.
- PAIM, J.S. A reorganização das práticas em distritos sanitários. In: MENDES, E.V. *Distrito sanitário. O processo social de mudança das práticas sanitárias do Sistema Único de Saúde*. São Paulo/Rio de Janeiro: HUCITEC/ABRASCO, 1993.
- PAIM, J.S. A reforma sanitária e os modelos assistenciais. In: ROUQUAYROL; ALMEIDA-FILHO. *Epidemiologia e saúde*. Rio de Janeiro: MEDSI, 1999. p. 473-488.
- PAIM, J.S. *Saúde, política e reforma sanitária*. Salvador: CEPS-ISC, 2002.
- PAIM, J.S. Políticas de saúde no Brasil. In: ROUQUAYROL; ALMEIDA-FILHO. *Epidemiologia e saúde*. Rio de Janeiro: MEDSI, 2003, p. 600. (no prelo).
- PATTON, M.Q. *How to use qualitative methods in evaluation*. Newbury Park London-New Delhi: SAGE publications, 1987.
- PATTON, M.Q. *Utilization-focused evaluation. The new century text*. Thousands Oaks- London- New Delhi: SAGE Publications, 1997.
- PATTON, M.Q. "Evaluation, knowledge management, best practices, and high quality lessons learned." *American Journal of Evaluation*, v. 22, n. 3, p. 329-336, 2001.
- PATTON, M.Q. *Qualitative research and evaluation methods*. Thousands Oaks London- New Delhi: SAGE Publications, 2002.
- PAWSON, R.; TILLEY, N. "Realistic evaluation bloodlines." *American Journal of Evaluation*, v. 22, n. 3, p. 317-324, 2001.
- PERELMAN, C. *Ética e direito*. Tradução: Maria Ermantina Galvão G. Pereira. São Paulo: Martins Fontes, 1996.
- ROGERS, P.J. "The whole world is evaluating half-full glasses." *American Journal of Evaluation*, v. 22, n. 3, p. 431-435, 2001.

- ROSSI, P.H.; FREEMAN, H.E. *Evaluation, a systematic approach*. Beverly Hills: Sage Publications, 2004.
- SCRIVEN, M. *Evaluation thesaurus*. Thousands Oaks London- New Delhi: SAGE Publications, 1991.
- SCRIVEN, M. "Evaluation: future tense." *American Journal of Evaluation*, v. 22, n. 3, p. 301-307, 2001.
- SMITH, M.F. "Evaluation: preview of the future." *American Journal of Evaluation*, v. 22, n. 3, p. 281-300, 2001.
- SMITH, N. "The feasibility and desirability of experimental methods in evaluation." *Evaluation and Program Plannin*, n. 3, p. 251-256, 1980.
- SOBERÓN, G. "La extensión de cobertura de los servicios de salud." *Gaceta Médica de México*, v. 124, n. 5-6, p. 163-166, 1988.
- STECHEER, B.M.; DAVIS, W.A. *How to focus an evaluation*. Newbury Park, London, New Delhi: Sage Publications, 1987.
- SUCHMAN, E.A. *Evaluative Research- Principles and Practice in Public Service and Social Action Programs*. New York: Russel SAGE Foundation, 1967.
- TRAVASSOS, C.; VIACAVAL, F.; PINHEIRO, R.; BRITO, A. "Utilization of health care services in Brazil: gender, family characteristics and social status." *Revista Panamericana de Salud Pública*, v. 11, n. 5-6, p. 365-73, 2002.
- VIEIRA-DA-SILVA, L.M. "Avaliação do processo de descentralização das ações de saúde." *Ciência e Saúde Coletiva*, v. 4, n. 2, p. 331-339, 1999a.
- VIEIRA-DA-SILVA, L.M. "The field of evaluation and the sur mesure strategy." *Cadernos de Saúde Pública*, v. 15, n. 2, p. 247-248, 1999b.
- VIEIRA-DA-SILVA, L.M.; FORMIGLI, V.; CERQUEIRA, M.; KRUCHEVSKY, L. "Coberturas vacinais superestimadas? Novas evidências a partir do inquérito de Pau da Lima." *Revista Panamericana de Salud Pública*, v. 1, p. 444-450, 1997.
- VIEIRA-DA-SILVA, L.M.; FORMIGLI, V.L.A. "Avaliação em saúde: limites e perspectivas." *Cadernos de Saúde Pública*, v. 10, n. 1, p. 80-91, 1994.
- VIEIRA-DA-SILVA, L.M.; FORMIGLI, V.L.A.; CERQUEIRA, M.P.; KRUCHEVSKY, L.; TEIXEIRA, M.M.A.; BARBOSA, A.S.M.; CONCEIÇÃO, P.S.A.; KHOURI, M.A.; NASCIMENTO, C.L. "O processo de distritalização e a utilização de serviços de saúde - avaliação do caso de Pau da Lima." *Cadernos de Saúde Pública*, v. 11, n. 1, p. 72 - 84, 1995.
- VUORI, H.V. *Quality Assurance of Health Services. Concepts and Methodology*. Copenhagen, WHO, Regional Office for Europe, 1982.
- WHO. *Health Impact Assessment: main concepts and suggested approach*. Gothenburg consensus paper. Gothenburg, WHO- European Centre for Health Policy-ECHP, 1999.
- WHO. *World Health Report*. Geneva, World Health Organization, 2000.
- YIN, R. K. *Case study research. Design and methods*. Beverly Hills, California: Sage Publications, 1994.