

Contextualizando a influência das avaliações: um ensaio meta-avaliativo com ênfase na utilidade e responsabilidade

Context to the influence of evaluations:

An essay on meta-evaluation with usefulness and accountability enfasis

Deborah Carvalho Malta

Departamento Materno infantil e de Saúde Pública, Escola de Enfermagem, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

Alanna Gomes da Silva

Programa de Pós-Graduação da Escola de Enfermagem, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil

Cátia Sá Guerreiro

Universidade Nova de Lisboa (UNL), Instituto de Higiene e Medicina Tropical (IHMT), Global Health and Tropical Medicine (GHTM) – Lisboa, Portugal

Zulmira M. A. Hartz

Professora Catedrática Convidada, GHTM, Instituto de Higiene e Medicina Tropical, Universidade NOVA de Lisboa, Portugal

Resumo

Objetivo: sistematizar a metodologia da meta-avaliação em seus aspectos conceituais e operacionais para facilitar o seu uso nas práticas de gestão em saúde.

Métodos: trata-se de um ensaio sobre a meta-avaliação. As bases teóricas foram pesquisadas em livros e artigos, publicados entre 1991 e 2018, que abordassem os aspectos conceituais, metodológicos e operacionais da meta-avaliação.

Resultados: a meta-avaliação consiste em um processo de delineamento, obtenção e aplicação da informação descritiva e de julgamento sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação em sua natureza sistemática, fundamentada em padrões e critérios propostos e validados. Existem diversas formas ou abordagens para realização da meta-avaliação. No entanto, as normas comumente seguem as diretrizes da *Joint Committee on Standards for Educational Evaluation*, que propõe 30 padrões agrupadas em cinco princípios: utilidade, viabilidade, propriedade, acurácia/precisão e responsabilidade, sendo esses essenciais e determinantes para qualidade de um estudo avaliativo.

Conclusão: a meta-avaliação contribui para o julgamento da qualidade dos processos avaliativos, na adequação e utilidade, podendo ser útil, em especial, quando se trata de temas e contextos que já foram avaliados previamente, e, por isso, também é descrita como “uma avaliação de uma avaliação”.

Palavras-chave:

Meta-avaliação, avaliação em saúde, avaliação de programas e projetos de saúde, gestão em saúde, melhoria de qualidade.

Abstract

Objective: to systematize the meta-evaluation methodology in its conceptual and operational aspects to facilitate its use in health management practices.

Methods: this is an essay on meta-evaluation. The theoretical bases were searched in books and articles, published between 1991 and 2018, that approached the conceptual, methodological and operational aspects of the meta-evaluation.

Results: the meta-evaluation consists of a process of delineating, obtaining and applying descriptive information and judgment about the usefulness, feasibility, adequacy and precision of an evaluation in its systematic nature, based on standards and criteria proposed and validated. There are several ways or approaches to the meta-evaluation. However, the standards commonly follow the guidelines of the *Joint Committee on Standards for Educational Evaluation*, which proposes 30 standards grouped into five principles: utility, viability, ownership, accuracy / precision and accountability, being these essential and determinants for the quality of an evaluative study.

Conclusion: the meta-evaluation contributes to the judgment of the quality of the evaluation processes, in the adequacy and usefulness, and can be useful, especially when it comes to themes and contexts that have already been evaluated and therefore is also described as "an evaluation of an evaluation".

Key words:

Meta-evaluation, health evaluation, program evaluation, health management, quality improvement.

Introdução

A avaliação de serviços de saúde tem sido cada vez mais aplicada em diferentes países, em função da sua utilidade e capacidade de apoiar os processos decisórios [1]. A avaliação consiste na determinação sistemática da qualidade ou valor de algum programa, serviço ou plano de ação, sendo utilizada na identificação de problemas, para corrigir erros e introduzir mudanças [1].

Avaliar implica na emissão de juízos, no estabelecimento de critérios, na precisão do objeto de estudo e diagnóstico. Além de gerar sugestões para tomada de decisão, se configurando fundamental para a gestão [2]. A avaliação produz conclusões válidas, e, por isso, deve ser conduzida de maneira ética, oportuna e com credibilidade, além de ser econômica e relevante [3,4].

O aumento dos processos avaliativos ocorreu devido a ampliação do campo, dos grupos de pesquisa na temática e dos avanços metodológicos. Entretanto, torna-se um desafio verificar se o conjunto de avaliações tem cumprido o seu papel de apoio à gestão e contribuído para o aperfeiçoamento das práticas [5].

A metodologia adequada para verificar a qualidade dos processos avaliativos consiste na meta-avaliação [5,6], que promove a análise sistemática dos estudos avaliativos, verificando se os objetivos foram atingidos e os resultados pretendidos alcançados [2,5]. A meta-avaliação produz a reflexão sobre todos os procedimentos avaliativos, definindo sua qualidade e oportunidade [7], além de apoiar os processos decisórios [8]. Entretanto, a condução da meta-avaliação ainda é restrita e torna-se necessária avançar na sua aplicação e compreensão.

Neste sentido, este estudo tem como objetivo sistematizar a metodologia da meta-avaliação em seus aspectos conceituais e operacionais para facilitar o seu uso nas práticas de gestão em saúde.

Métodos

Trata-se de um ensaio sobre a meta-avaliação que, conforme Silveira (1992) [9], visa aprofundar o debate sobre o tema por meio da exposição das ideias e pontos de vista, buscando originalidade no enfoque, sem, contudo, explorar o tema de forma exaustiva. O ensaio consiste em um estudo formal, discursivo e concludente que expõe o tema de forma lógica e reflexiva. Assim, para a elaboração do presente trabalho, as bases teóricas e conceituais foram pesquisadas em livros e artigos, independentes do ano de publicação. Foram buscados artigos que abordassem a me-

ta-avaliação sobre os aspectos conceituais, metodológicos e operacionais, visando apoiar os gestores e avaliadores em futuras avaliações. Foram utilizados estudos publicados nos idiomas inglês, português ou espanhol. Foram excluídos textos completos não disponíveis e trabalhos acadêmicos como monografias, dissertações e teses. O período de publicação dos periódicos consultados foi entre os anos 1991 e 2018.

Resultados

Conceito

A Joint Committee on Standards for Educational Evaluation (JCSEE) (2011) [10], define a meta-avaliação como o processo de delineamento, obtenção e aplicação da informação descritiva e de julgamento sobre a utilidade, a viabilidade, a adequação e a precisão de uma avaliação. Deve ser feita com integridade, honestidade, respeitabilidade e responsabilidade social. Para, assim, orientar a avaliação e divulgar publicamente seus pontos fortes e fracos.

A meta-avaliação pode ser descrita como “uma avaliação de uma avaliação” [6,11], sendo revisões sistemáticas de avaliações que ajudam a determinar a qualidade dos processos, serviços e resultados das avaliações [12].

Operacionalmente a meta-avaliação é fundamentada em padrões e critérios propostos e validados pelas associações profissionais e outros órgãos auditores. Ademais, busca-se a verificação da coerência teórica e prática da pesquisa com os padrões internacionais e governamentais do controle de qualidade das avaliações de políticas públicas. [8,10,13] As meta-avaliações são necessárias em todos os tipos de avaliação, incluindo programas, projetos, produtos, sistemas, instituições, teorias, modelos e pessoal [6].

A meta-avaliação também é importante pela oportunidade de incorporar novos conhecimentos sobre os resultados alcançados, além de ser um método para a certificação da qualidade [4,14].

Operacionalização da meta-avaliação

A operacionalização da meta-avaliação deve ter rigor na aplicação, pois exigirá um conjunto de procedimentos, padrões e critérios para o seu julgamento, o que significa emitir uma nova avaliação sobre o estudo avaliativo [4,8]. Nesse sentido, alguns passos metodológicos devem ser seguidos para facilitar a compreensão do processo. A seguir, alguns exemplos dos métodos para realização da meta-avaliação.

Passo 1 - Planejar a meta-avaliação

A primeira etapa consiste em realizar de forma criteriosa o planejamento da meta-avaliação, o que inclui a busca e organização do material a ser avaliado, a escolha da equipe que conduzirá, a capacitação da equipe de pesquisadores, a definição das estratégias e objetivos da avaliação.

Outra etapa fundamental consiste no mapeamento e na definição dos atores envolvidos. Quem são os interessados e envolvidos nesta avaliação? Quais e quem são os atores-chaves? Após a definição dos envolvidos na avaliação, deve-se realizar junto a eles o levantamento de demandas e das perguntas a serem respondidas. Essas questões são essenciais para se definir os recursos e tempo necessário para a condução do processo.

Passo 2 - levantamento do material a ser meta-avaliado

Revisão da literatura ou busca de relatórios

Devem ser definidos de forma sistemática os procedimentos para a revisão da literatura tais como: formulação da questão norteadora da pesquisa; definição dos descritores, critérios de elegibilidade (inclusão e exclusão de estudos) e bases de dados; elaboração da estratégia de pesquisa, bem como os procedimentos de seleção dos estudos, extração e sumarização dos dados. Esses procedimentos sistemáticos permitirão coletar, identificar, selecionar, sintetizar as evidências de pesquisas, para mapear a literatura existente sobre a meta-avaliação em termos de natureza, características e volume [15].

As meta-avaliações também podem ser conduzidas não por material publicado, tais como relatórios, que reúnem avaliações prévias conduzidas [8].

Espera-se que a meta-avaliação especifique seu referencial teórico para permitir que os diversos públicos e avaliadores sigam os passos e que possam discutir suas descobertas e julgamentos [16].

Após a pesquisa da literatura, deve-se pré-selecionar os estudos e avaliá-los, para identificar com mais precisão a sua relevância e se os critérios de inclusão e exclusão estão contemplados. Após essa etapa, os dados/informações são extraídos. Para facilitar a extração dos dados pode-se utilizar formulários, os quais irão conter a síntese dos estudos, como títulos, autores, nome do programa avaliado, ano de avaliação, tipo de avaliação, se foi utilizado avaliador interno ou externo, amostra, metodologias empregadas, tipo de estudo, validade

interna e externa, fonte de financiamento, dentre outros [17]. O preenchimento do formulário, representa a primeira aproximação com o objeto, dimensiona os estudos e relatórios que serão empregados na meta-avaliação. Portanto, o formulário deve ser sensível para capturar a natureza das avaliações e o mérito técnico dos estudos de avaliação.

Passo 3 – Avaliar a qualidade dos estudos (seleção dos padrões)

Definir os principais padrões

Com os estudos já selecionados é essencial verificar a qualidade das avaliações, por meio da utilização de padrões ou normas avaliativas, tendo geralmente como base os parâmetros utilizados pelo sistema educacional americano, que orientada as agências de avaliação em todo o mundo [10]. Os padrões apoiam o julgamento de cada avaliação, além de padronizar os processos, possibilitando comparações e conclusões [4]. Em 1981, a JCSEE publicou a primeira edição dos padrões para avaliação de programas, projetos e materiais educacionais [18]. Em 1994, a segunda edição desses padrões foi atualizada [19]. Em 2011, após amplo processo de consulta, envolvendo especialistas de diferentes países, com a tarefa de investigar e revisar os padrões de avaliação, foi publicado a terceira edição [10], a qual oferece padrões para a execução de avaliações e das meta-avaliações, sendo amplamente adotados em diversos países [4,8].

Existem várias formas ou abordagens para realização da meta-avaliação e que podem ser úteis aos avaliadores e interessados na qualidade das avaliações realizadas [4]. No entanto, as normas para meta-avaliação comumente seguem as diretrizes da JCSEE, que propõe 30 padrões (check list) que são agrupadas em cinco categorias ou princípios: utilidade, viabilidade, propriedade, acurácia/precisão e responsabilidade, sendo essenciais e determinantes para qualidade de um estudo avaliativo [20].

Nesse sentido, o princípio da utilidade visa atender as necessidades dos grupos de interesse. A viabilidade, destina-se a aumentar a eficácia e a eficiência da avaliação. A propriedade relaciona-se com o que é adequado, justo, legal, correto e ético nas avaliações, além do respeito pelos envolvidos. Os princípios de acurácia/precisão aumentam a confiabilidade e a veracidade das representações, proposições e descobertas da avaliação, principalmente aquelas com interpretações e julgamentos sobre qualidade. A avaliação de responsabilidade encoraja a documentação adequada das avaliações e uma perspectiva

meta-avaliativa focada na melhoria e na responsabilização por processos e produtos de avaliação [8,20]. A seguir os cinco padrões são apresentados com mais detalhes e todos estão em consonância com a Joint Committee on Standards for Educational Evaluation, 2018 [20]:

1) Utilidade: totalizam oito padrões, os quais procuram responder as questões: A avaliação foi útil? Atendeu as demandas dos usuários, interessados na avaliação? Esses padrões divididos em: U1 Credibilidade do avaliador: as avaliações devem ser conduzidas por pessoas qualificadas para manter a credibilidade no contexto da avaliação; U2 Atenção aos grupos de interesse (stakeholders): As avaliações devem dedicar atenção a toda a gama de indivíduos e grupos inseridos no programa e afetados por sua avaliação; U3 Propostas negociadas: As finalidades da avaliação devem ser identificadas e continuamente negociadas com base nas necessidades dos stakeholders; U4 Valores Explícitos: As avaliações devem esclarecer e especificar os valores individuais e culturais que sustentam propósitos, processos e julgamentos; U5 Informações Relevantes: As informações de avaliação devem atender às necessidades identificadas e emergentes dos stakeholders; U6 Processos e Produtos Significativos: As avaliações devem construir atividades, descrições e julgamentos de forma a encorajar os participantes a redescobrir, reinterpretar ou revisar seus entendimentos e comportamentos; U7 Comunicações e relatórios oportunos e apropriados: as avaliações devem atender às necessidades contínuas de informação de seus múltiplos públicos; U8 Preocupação com as consequências e a influência: As avaliações devem promover o uso responsável e adaptativo, ao mesmo tempo em que se protege contra consequências negativas não intencionais e uso indevido.

2) Viabilidade, exequibilidade ou factibilidade: são quatro padrões que se destinam a aumentar a eficácia e a eficiência da avaliação e se referem aos efeitos de contextos, culturas, custos, política, poder e recursos disponíveis, sendo: F1 Gerenciamento de projetos: as avaliações devem usar estratégias eficazes de gerenciamento de projetos; F2 Procedimentos práticos: Os procedimentos de avaliação devem ser práticos e receptivos à maneira como o programa opera; F3 Viabilidade Contextual: As avaliações devem reconhecer, monitorar e equilibrar os interesses e necessidades culturais e políticas de indivíduos e grupos; F4 Uso de recursos: As avaliações devem usar os recursos de forma eficaz e eficiente.

3) Propriedade ou adequação: são sete padrões

que asseguram que a avaliação foi desenvolvida respeitando aspectos morais, éticos e legais concernentes à qualidade da avaliação, sendo: P1 Orientação Responsiva e Inclusiva: As avaliações devem ser responsivas aos stakeholders e suas comunidades; P2 Acordos Formais: Acordos de avaliação devem ser negociados para tornar as obrigações explícitas e levar em conta as necessidades, expectativas e contextos culturais dos clientes e outros stakeholders; P3 Direitos humanos e respeito: As avaliações devem ser planejadas e conduzidas para proteger os direitos humanos e legais e manter a dignidade dos participantes e de outros stakeholders; P4 Clareza e justiça: As avaliações devem ser compreensíveis e justas para atender às necessidades e objetivos dos stakeholders; P5 Transparência e Divulgação: As avaliações devem fornecer descrições completas das conclusões, limitações e conclusões a todas os stakeholders, a menos que isso viole as obrigações legais e de propriedade; P6 Conflitos de Interesses: As avaliações devem identificar de forma aberta e honesta e abordar os conflitos de interesses reais ou percebidos que possam comprometer a avaliação; P7 Responsabilidade Fiscal: As avaliações devem contabilizar todos os recursos gastos e obedecer a procedimentos fiscais sólidos.

4) Acurácia ou precisão: são oito padrões que destinam-se a aumentar a confiabilidade e a veracidade das representações, proposições e descobertas de avaliação, especialmente aquelas que apóiam interpretações e julgamentos sobre qualidade, sendo: A1 Conclusões e decisões justificadas: As conclusões e decisões de avaliação devem ser explicitamente justificadas nas culturas e contextos em que têm consequências; A2 Informações válidas: As informações da avaliação devem atender aos objetivos pretendidos e apoiar interpretações válidas; A3 Informações confiáveis: Os procedimentos de avaliação devem fornecer informações suficientemente confiáveis e consistentes para os usos pretendidos; A4 Programa Explícito e Descrições de Contexto: As avaliações devem documentar os programas e seus contextos com detalhes e escopo apropriados para fins de avaliação; A5 Gerenciamento de Informações: As avaliações devem empregar métodos sistemáticos de coleta, revisão, verificação e armazenamento de informações; A6 Desenhos e Análises de Som: As avaliações devem empregar desenhos e análises tecnicamente adequados que sejam apropriados para os propósitos de avaliação; A7 Explícito Raciocínio de avaliação: O raciocínio de avaliação leva de informações e análises às descobertas, interpretações, conclusões e julgamentos deve ser documentado de forma clara e completa; A8

Comunicação e Relatórios: As comunicações de avaliação devem ter escopo adequado e proteção contra equívocos, desvios, distorções e erros.

5) Responsabilidade: totalizam três padrões que incentivam a documentação adequada das avaliações e uma perspectiva meta-avaliativa focada na melhoria e responsabilização por processos e produtos de avaliação, sendo: E1 Documentação de avaliação: As avaliações devem documentar completamente seus objetivos negociados e os projetos, procedimentos, dados e resultados implementados; E2 Meta-avaliação interna: Os avaliadores devem usar esses e outros padrões aplicáveis para examinar a responsabilidade do projeto de avaliação, procedimentos empregados, informações coletadas e resultados; E3 Meta-avaliação externa: Os patrocinadores da avaliação do programa, clientes, avaliadores e outros stakeholders devem incentivar a realização de meta-avaliações externas usando esses e outros padrões aplicáveis.

Os padrões de avaliação devem ser selecionados pelo grupo de avaliadores, buscando consenso sobre os padrões mais adequados à meta-avaliação em curso. Os procedimentos para análise de cada avaliador devem ser realizados de forma independente e após faz-se as comparações e classificações entre os avaliadores, verificando assim a concordância e o consenso quanto aos padrões avaliados. Nesse sentido, negociações e tomada de decisões são necessárias na adaptação dos padrões por cada meta-avaliador. Internacionalmente, tem-se discutido sobre a necessidade de “padrões abertos” devido às dificuldades inerentes à transferência de parâmetros entre diferentes culturas e origens, bem como pela complexidade das políticas e programas que são avaliadas. Contudo, o desenvolvimento desses padrões específicos ainda não é uma realidade [16].

Analisar os padrões

Trata-se de um procedimento complexo, o qual consiste na atribuição de valor a cada padrão, podendo estar impregnado de subjetividade, caso não sejam aplicadas técnicas adequadas. Assim, para reduzir a subjetividade, o avaliador deve buscar todas as informações inerentes ao objeto. Quando a meta-avaliação for realizada mediante revisão da literatura, os artigos devem ser avaliados com todas as informações disponíveis [21]. Caso seja possível, incluir outras fontes de dados, bases de dados e relatórios da pesquisa. No caso de meta-avaliação de planos e programas, deve-se incluir todos os documentos e/ou relatórios disponíveis e outras informações complementares, junto

aos participantes ou envolvidos no processo avaliativo por meio de entrevistas, grupos focais [8]. Também é desejável que seja ampliado o processo de busca e de coleta de dados, visando ter-se uma ampla visão do processo.

Análise de conteúdo

Para análise de conteúdo dos estudos, pode-se também utilizar a metodologia de Bardin [22], o qual designa o termo “análise de conteúdo” como um conjunto de técnicas de análise das comunicações, visando obter por procedimentos sistemáticos e objetivos a descrição do conteúdo das mensagens e indicadores que permitam a inferência de conhecimentos relativos às condições de produção/recepção [22]. Assim, os conteúdos são identificados e classificados, definindo as categorias, mediante o confronto das entrevistas, com o material documental e da análise de contexto.

Seleção dos padrões

Torna-se importante ressaltar que a escolha dos padrões estará relacionada com os conteúdos encontrados nos estudos, relatórios e dos programas. Após a seleção dos artigos e análises dos conteúdos/resultados, serão verificados quais dos 30 padrões poderão ser aplicados na meta-avaliação. O padrão E2 já identifica que se deve “Designar ou definir os padrões a serem usados no julgamento da avaliação”. Por isso, as meta-avaliações terão variações dos padrões, visto que a escolha será conduzida segundo a análise de conteúdo dos resultados encontrados previamente.

Pontuação ou escores

Para cada um dos padrões, os *checkpoints* mais importantes são indicados. Sugere-se que cada padrão seja pontuado em cada *checkpoint*, sendo feito da seguinte forma: Excelente (9-10); Muito bom (7-8); Bom (5-6); Fraco (4-3) e Crítico (0-2). Considera-se que uma avaliação seja falha se pontuar em Crítico (0-2) nas normas: P1 Orientação Responsiva e Inclusiva, A1 Conclusões e decisões justificadas, A2 Informações válidas e A8 Comunicação e Relatórios [24]. O quadro 1 apresenta os 30 padrões e seus respectivos *checkpoints* e pontuações.

O meta-avaliador deverá assinalar seu julgamento relativo a cada padrão na lista de verificação, após analisar o conteúdo do relatório de avaliação. Se vários avaliadores julgarem um mesmo relatório, os resultados assinalados devem ser confrontados e as discrepâncias devem ser discutidas para se chegar a um consenso [4].

Procede-se depois com a requalificação por princípio de avaliação, como exemplo, pode-se classificar em: Insatisfatório (<33,3%), Aceitável (33,3%-66,6%) e Satisfatório (>66,6%) [8].

Quadro 1 - Matriz das medidas e fonte de recolha de dados

Para atender o padrão de UTILIDADE, o programa de avaliação deverá:
<p>U1 Credibilidade do avaliador</p> <ul style="list-style-type: none"> () Envolver o avaliador em quem os stakeholders confiam () Envolver avaliadores que sejam sensíveis as questões de gênero, status socioeconômico, raça e as diferenças linguísticas e culturais () Garantir que o plano de avaliação responda às principais preocupações dos stakeholders () Atender adequadamente às críticas e sugestões dos stakeholders () Manter os stakeholders informados sobre o progresso da avaliação.
<p>U2 Atenção aos stakeholders</p> <ul style="list-style-type: none"> () Identificar claramente o cliente de avaliação () Consultar os stakeholders em potencial para identificar suas necessidades de informação () Com o cliente, classifique os stakeholders por importância relativa () Envolver os stakeholders ao longo da avaliação () Abordar as necessidades de avaliação dos stakeholders
<p>U3 Propostas negociadas</p> <ul style="list-style-type: none"> () Monitorar e descrever como os objetivos da avaliação permaneceram os mesmos ou mudaram ao longo do tempo () Atualizar os procedimentos de avaliação para adaptar as mudanças nas finalidades da avaliação () Registrar os procedimentos de avaliação, conforme implementado () Descrever as finalidades e procedimentos da avaliação nos relatórios de avaliação no resumo e completa
<p>U4 Valores Explícitos</p> <ul style="list-style-type: none"> () Considerar fontes alternativas para interpretar os resultados da avaliação () Referenciar leis pertinentes () Referenciar a missão institucional () Referenciar dos objetivos do programa () Apresentar as interpretações alternativas baseadas nos conflitos, mas que tenham bases confiáveis
<p>U5 Informações Relevantes</p> <ul style="list-style-type: none"> () Entrevistar as partes interessadas para determinar suas diferentes perspectivas () Atribuir prioridade às questões mais importantes () Permitir flexibilidade para adicionar perguntas durante a avaliação () Obter informações suficientes para abordar as questões de avaliação mais importantes das partes interessadas () Obter informações suficientes para avaliar o mérito e o valor do programa
<p>U6 Processos e Produtos Significativos</p> <ul style="list-style-type: none"> () Relatar claramente as informações essenciais () Emitir relatórios breves, simples e diretos () Relatórios com foco nas questões contratadas () Descrever os propósitos, procedimentos e conclusões da avaliação () Apoiar conclusões e recomendações () Fornecer um resumo executivo
<p>U7 Comunicações e relatórios oportunos e apropriados</p> <ul style="list-style-type: none"> () Fazer relatórios provisórios e oportunos para os usuários previstos () Entregar o relatório final quando for necessário () Ter trocas oportunas com a equipe do programa () Ter trocas oportunas com os clientes do programa () Manter as apresentações apropriadamente breves
<p>U8 Preocupação com as consequências e a influência</p> <ul style="list-style-type: none"> () Incentivar e apoiar o uso dos resultados pelas partes interessadas () Fornecer relatórios provisórios () Certifique-se de que os relatórios estejam abertos, verdadeiros e concretos () Completar relatórios escritos com comunicação oral em curso () Realize workshops de feedback para examinar e aplicar os resultados
<p>() 9-10 Excelente () 7-8 Muito bom () 5-6 Bom () 4-3 Fraco () 0-2 Crítico</p>

Para atender o padrão de VIABILIDADE, o programa de avaliação deverá:**F1 Gerenciamento de projetos**

- Relacionar-se de forma consistente com todas as partes interessadas de forma profissional
- Minimizar a interrupção
- Garantir os direitos de privacidade dos participantes
- Responder as preocupações dos participantes sobre a avaliação
- Não ignore ou ajude encobrir a incompetência, comportamento antiético, fraude, desperdício ou abuso de qualquer participante

F2 Procedimentos práticos

- Adaptar métodos e instrumentos aos requisitos de informação
- Minimizar a carga de dados
- Nomear pessoas competente
- Escolha procedimentos de acordo com as limitações existentes
- Faça um cronograma factível
- Envolver os moradores locais para ajudar a realizar a avaliação

F3 Viabilidade Contextual

- Evitar ou neutralizar tentativas vieses ou aplicar mal os resultados
- Concordar com a autoridade editorial e de divulgação
- Comunicar as visões divergentes

F4 Uso de recursos

- Seja eficiente
- Informe as decisões
- Promova a melhoria do programa
- Gere novos pensamentos
- Minimizar as demandas de tempo do pessoal do programa

9-10 Excelente 7-8 Muito bom 5-6 Bom 4-3 Fraco 0-2 Crítico

Para atender o padrão de PROPRIEDADE, o programa de avaliação deverá:**P1 Orientação Responsiva e Inclusiva**

- Avalie os resultados do programa em relação às necessidades apresentadas pelos clientes
- Faça a orientação da avaliação de forma clara para os stakeholders
- Identifique os pontos fortes para construir o programa
- Identifique os pontos fracos para corrigir o programa

P2 Acordos Formais

- Objetivo de avaliação e perguntas
- Relatórios de avaliação
- Lançamento de relatórios
- Procedimentos de avaliação e cronograma
- Confidencialidade / anonimato dos dados
- Recursos de avaliação

P3 Direitos humanos e respeito

- Esclarecer os usos pretendidos da avaliação
- Mantenha os stakeholders informados
- Respeite a diversidade
- Honrar os acordos de confidencialidade / anonimato
- Não faça nenhum mal

P4 Clareza e justiça

- Avalie e relate os pontos fortes do programa
- Avalie e relate os pontos fracos do programa
- Relatório sobre os resultados pretendidos
- Relatório sobre resultados não intencionais
- Apropriadamente endereçar críticas do projeto de relatório
- Estimar e relatar os efeitos das limitações da avaliação no julgamento geral do programa

9-10 Excelente 7-8 Muito bom 5-6 Bom 4-3 Fraco 0-2 Crítico

P5 Transparência e Divulgação

- Relatar todas as descobertas por escrito
- Relatar os pontos de vista relevantes de ambos os apoiadores e críticos do programa
- Relatar conclusões e recomendações equilibradas e informadas
- Mostrar a base para as conclusões e recomendações
- Divulgue as limitações da avaliação

P6 Conflitos de Interesses

- Identificar potenciais conflitos de interesse no início da avaliação
- Fornecer salvaguardas escritas e contratuais contra conflitos de interesse identificados
- Envolver vários avaliadores
- Manter registros de avaliação para revisão independente
- Quando apropriado, divulgar publicamente os procedimentos de avaliação, dados e relatórios

P7 Responsabilidade Fiscal

- Especifique antecipadamente o orçamento para as despesas
- Manter o orçamento suficientemente flexível para permitir realocações apropriadas para melhorar a avaliação
- Obter aprovação apropriada para as modificações orçamentárias necessárias
- Manter registros precisos das fontes de financiamento e despesas
- Manter registros sobre alocações de trabalho e tempo gasto no trabalho
- Seja cauteloso em gastar recursos de avaliação

9-10 Excelente 7-8 Muito bom 5-6 Bom 4-3 Fraco 0-2 Crítico

Para atender o padrão de ACURÁCIA, o programa de avaliação deverá:

A1 Conclusões e decisões justificadas

- Concentre as conclusões diretamente nas questões de avaliação
- Refletir com precisão os procedimentos e conclusões da avaliação
- Limitar as conclusões aos períodos de tempo, contextos, finalidades e atividades aplicáveis
- Cite a informação relacionadas a cada conclusão
- Relatar explicações alternativas plausíveis dos resultados
- Explique o que foi rejeitado

A2 Informações válidas

- Enfoque a avaliação nas principais questões
- Conforme apropriado, empregue várias medidas para abordar cada questão
- Treine e capacite aqueles que irão coletar os dados
- Documentar e relatar as condições e processos de coleta de dados
- Documentar como as informações de cada procedimento foram pontuadas, analisadas e interpretadas
- Avaliar e relatar a abrangência das informações fornecidas pelos procedimentos, a relação das informações necessárias para responder ao conjunto de questões da avaliação

A3 Informações confiáveis

- Relatar a confiabilidade do instrumento e avaliar os fatores que influenciaram a confiabilidade, incluindo as características dos examinandos, as condições de coleta de dados e os vieses do avaliador
- Treine e capacite os pontuadores e analistas para produzir resultados consistentes
- Teste novos instrumentos para identificar e controlar fontes de erro
- Conforme apropriado, envolva e verifique a consistência entre os observadores
- Reconheça problemas de confiabilidade no relatório final

A4 Programa Explícito e Descrições de Contexto

- Recolha as descrições do programa pretendido pelo cliente e stakeholders
- Manter registros de como o programa foi produzido
- Se possível, envolva observadores independentes para descrever as operações reais do programa
- Descreva como o programa realmente funcionou
- Analise as discrepâncias entre como o programa foi planejado para operar e como ele realmente operou

A5 Gerenciamento de Informações

- Obter informações de várias fontes
- Conforme apropriado, empregue vários métodos de coleta de dados
- Documentar, justificar e relatar os critérios e métodos usados para selecionar as fontes de informação
- Documentar, justificar e relatar os meios usados para obter informações de cada fonte
- Incluir instrumentos de coleta de dados em um apêndice técnico ao relatório de avaliação
- Sistematizar e controlar o armazenamento da informação de avaliação

<p>A6 Desenhos e Análises de Som</p> <ul style="list-style-type: none"> () Estabelecer protocolos para controle de qualidade da informação de avaliação () Treinar a equipe de avaliação para aderir aos protocolos () Verificar sistematicamente a precisão da pontuação e codificação () Quando viável, use vários avaliadores e verifique a consistência de seu trabalho () Escolha procedimentos apropriados para as questões de avaliação e natureza do período () Obter informações chaves para as importantes questões de avaliação
<p>A7 Explícito Raciocínio de avaliação</p> <ul style="list-style-type: none"> () Empregar vários procedimentos analíticos para verificar a consistência e replicabilidade dos resultados () Examine a variabilidade, bem como tendências centrais () Identifique e examine <i>outliers</i> e sua exatidão () Use <i>displays</i> visuais para esclarecer a apresentação e interpretação dos resultados estatísticos () Tire conclusões e recomendações e demonstre sua importância
<p>A8 Comunicação e Relatórios</p> <ul style="list-style-type: none"> () Relatórios de garantia de distorções deliberadas ou inadvertidas () Relate as perspectivas de todos os stakeholders () Relate as plausíveis conclusões alternativas () Descreva os passos para controle do viés () Participe de apresentações públicas dos resultados para ajudar a proteger e corrigir distorções de outras partes interessantes
<p>() 9-10 Excelente () 7-8 Muito bom () 5-6 Bom () 4-3 Fraco () 0-2 Crítico</p>
<p>Para atender o padrão de RESPONSABILIDADE, o programa de avaliação deverá:</p>
<p>E1 Documentação de avaliação</p> <ul style="list-style-type: none"> () Coletar descrições dos projetos de avaliação implementados () Coletar descrições dos procedimentos de avaliação () Registrar todos os dados coletados () Analisar os dados e registrar os resultados () Produzir um relatório técnico que forneça informações sobre o desenho da avaliação, procedimentos, dados e resultados
<p>E2 Meta-avaliação interna</p> <ul style="list-style-type: none"> () Designar ou definir os padrões a serem usados no julgamento da avaliação () Atribuir a alguém a responsabilidade de documentar e avaliar o processo e os produtos de avaliação () Orçamento de forma adequada e suficiente para realizar a meta-avaliação
<p>E3 Meta-avaliação externa</p> <ul style="list-style-type: none"> () Orçamento de forma adequada e suficiente para realizar a meta-avaliação () Manter um registro de todas as etapas da meta-avaliação, informações e análises
<p>() 9-10 Excelente () 7-8 Muito bom () 5-6 Bom () 4-3 Fraco () 0-2 Crítico</p>

Fonte: Adaptado Stufflebeam, 2012 [24].

Passo 4 - Sintetizando as conclusões da avaliação

Dando sequência ao processo, os avaliadores devem revisar e comparar os escores dos padrões, buscando o consenso e a síntese final. Cada padrão definido será pontuado, obtendo-se o escore final. Assim, os achados devem ser sintetizados, buscando responder se as avaliações atenderam aos padrões previamente descritos e visando o julgamento da qualidade das avaliações. O resultado consiste em uma síntese sucinta da pesquisa [17].

Na síntese as seguintes perguntas devem ser respondidas: O que os resultados das avaliações sugerem? Quais as principais descobertas? O programa é útil? Após responder a essas questões, deve-se proceder a uma revisão narrativa dos achados, apontando os resultados e as principais conclusões das avaliações para responder junto aos participan-

tes da meta-avaliação, se o programa teve impacto positivo. Além de preparar as recomendações que são importantes para o aperfeiçoamento do programa. Deve-se ainda discutir os resultados obtidos com a equipe do programa, com o avaliador, com os grupos de interesse e preparar o plano de comunicação dos resultados [17].

Aplicação de padrões de meta-avaliação

A seguir são apresentados exemplos de meta-avaliações que foram conduzidas na avaliação do Plano Nacional de Desenvolvimento Sanitário I (PNDS I) [25] de República da Guiné Bissau (RGB) [8].

Sendo um Estado independente desde 1973, a RGB estabeleceu sua Política Nacional de Saúde em 1993 e

tem feito ao longo dos anos o exercício do Planejamento Estratégico em Saúde (PES).

Do primeiro grande exercício de PES resultou o PNDS I, estabelecido com horizonte temporal de 1998-2002 [26]. Porém, a situação político-militar de 1998 comprometeu a execução do referido plano, tendo este sido revisto para o horizonte temporal de 2003-2007, com o objetivo de permitir que as atividades inicialmente programadas pudessem ser implementadas [25,27].

Em finais de 2007, foi efetuada a avaliação da execução do PNDS I, a qual foi validada pelo Ministério da Saúde Pública (MINSAP) [25]. Com o objetivo de avaliar a qualidade da avaliação efetuada procedeu-se a uma meta-avaliação [8].

Optou-se por utilizar 13 dos 30 padrões de meta-avaliação JCSEE na análise do relatório de avaliação. No trabalho atual optou-se por apresentar dois exemplos do uso dos princípios da Utilidade e Factibilidade ou Viabilidade.

Foi aplicada uma escala quantitativa de 0-10 a cada um dos padrões, pela qual os valores 9 e 10 correspondem a “excelente”; 7 e 8 a “muito bom”; 5 e 6 a “bom”; 3 e 4 a “fraco”; 0 a 2 a “crítico”. Procedeu-se depois à requalificação por princípio de avaliação em “Insatisfatório”, “Aceitável” e “Satisfatório” de acordo com a avaliação percentual <33,3%, 33,3%-66,6% e >66,6% respetivamente [21, 28,29].

Para a atribuição da classificação por padrão e assim obter o valor por princípio de meta-avaliação foi efetuada

Quadro 2 - Princípio de Utilidade: avaliação por padrões

PADRÃO	VALOR DE CLASSIFICAÇÃO	JUSTIFICATIVA	OBSERVAÇÃO
Classificação	29 = 96,7%	Classificação final: <i>Satisfatório</i>	
<i>Credibilidade do Avaliador</i>	10	A coordenação da equipe de avaliadores detinha amplo conhecimento da realidade contextual e do processo de implantação do PNDS I.	
<i>Clareza dos relatórios</i>	10	O relatório descreve claramente o programa avaliado, incluindo o contexto, os propósitos, procedimentos e conclusões da avaliação.	Informação obtida apenas na análise documental dos relatórios de avaliação.
<i>Impacto da Avaliação</i>	9	A avaliação foi acompanhada pelos interessados localmente, membros do MINSAP. O posterior processo de planeamento do PNDS II “teve por base não apenas o relatório de avaliação da implementação do PNDS I como também uma consultoria externa da OMS Genebra (...)”. A coordenação da equipe de avaliadores do PNDS I “foi sendo consultada ao longo do processo (...)”.	Informações obtidas com recurso das entrevistas realizadas.

Quadro 3 - Princípio de Factibilidade ou Viabilidade: avaliação por padrões

PADRÃO	VALOR DE CLASSIFICAÇÃO	JUSTIFICATIVA	OBSERVAÇÃO
Classificação	16 = 80%	Classificação final: <i>Satisfatório</i>	
<i>Procedimentos práticos</i>	7	Embora muito claramente definidos no documento de relatório de avaliação, e, portanto, bem definidos numa fase de preparação da avaliação, os procedimentos demonstraram ser de difícil implantação no momento da avaliação no terreno. Foram elaborados em Lisboa, com base no documento de PNDS I, do qual se pretendia avaliar a implantação. Porém, ocorreria uma guerra civil e “ao chegar ao terreno, os dados para calcular os indicadores definidos tinham desaparecido”. Os procedimentos demonstraram não ser tão práticos como poderiam parecer por se encontrarem desajustados da realidade local no momento da avaliação.	Informações complementares obtidas com recurso das entrevistas realizadas.
<i>Viabilidade contextual</i>	9	“O pedido de avaliação vem do MINSAP e não de algum dos financiadores ou parceiros”, não se excluindo a existência de diferentes grupos de interesse no programa em avaliação. Processo avaliativo apoiado e motivado pelo MINSAP.	Informações obtidas com recurso das entrevistas realizadas

a análise de conteúdo [8] do relatório de avaliação disponível e do próprio PNDS. A informação obtida foi complementada com recurso a análise de conteúdo de entrevistas realizadas a avaliadores e outros atores chave em PES na RGB.

A título de exemplo, os quadros 2 e 3 descrevem a justificativa para a atribuição das classificações aos padrões definidos. O Quadro 2 apresenta o princípio de utilidade e a escolha dos padrões de Credibilidade do avaliador, Clareza dos relatórios, Impacto da avaliação, o valor da classificação, a justificativa da pontuação e observações na análise.

O Quadro 3 apresenta o princípio de Factibilidade a escolha dos padrões de Procedimentos práticos, Viabilidade contextual, o valor da classificação, a justificativa da pontuação e observações na análise.

Discussão

Este ensaio define a meta-avaliação, como a avaliação das avaliações, verificando se os estudos avaliativos empregados correspondem aos princípios e requisitos de uma avaliação, bem como a necessidade de informar aos interessados a utilidade ou valor da avaliação primária [5,6]. O texto apresenta as definições operacionais, os princípios e padrões da meta-avaliação, além de apresentar um exemplo prático de aplicação dos padrões de Utilidade e Factibilidade ou Viabilidade, visando ampliar o debate sobre o tema e facilitar a sua incorporação.

A meta-avaliação é, operacionalmente, entendida como um processo de descrição, julgamento e síntese de estudos ou quaisquer procedimentos de avaliação com o objetivo de assegurar a qualidade dos estudos avaliativos [8]. Assim, a meta-avaliação permite o acesso à informação sobre as limitações e as potencialidades de avaliações realizadas, aumentando sua credibilidade e permitindo às partes interessadas julgar e contextualizar os resultados obtidos [8,30]. Além disso, identifica-se os pontos fortes e fracos na prática de avaliação [21] e verifica se os pilares avaliativos foram utilizados: a pertinência, a justiça e o caráter participativo da avaliação [31].

Hartz (2008) [21], defende que o principal desafio de uma meta-avaliação é verificar até que ponto os estudos analisados respondem aos pré-requisitos de uma adequada avaliação (mérito) e a necessidade de informações dos interessados (valor). Portanto, a aferição da qualidade é o tema central da meta-avaliação para avançar no aprendizado e a melhorar as práticas [2, 32].

A meta-avaliação distingue-se da pesquisa avaliativa por incluir um caráter normativo, objetivando, não apenas avaliar os efeitos de uma intervenção, e entender como foram obtidos, mas também julgá-la, visando seu aprimoramento [21].

Os avaliadores precisam de meta-avaliações para melhorar os estudos individuais, desenvolver novas abordagens de avaliação, ganhar e manter a credibilidade de seus serviços entre os clientes e outros avaliadores. Os consumidores/clientes precisam de meta-avaliações para evitar avaliações inválidas e obter informações confiáveis a partir das meta-avaliações, bem como para ajudar a assegurar que os serviços de avaliação de sua instituição sejam defensáveis, funcionais e que valham o investimento [6].

Segundo o tempo de realização, as meta-avaliações podem ser classificadas em formativas e somativas. As formativas são realizadas durante o processo de desenvolvimento da avaliação, com o objetivo de auxiliar os avaliadores a planejar, conduzir, melhorar, interpretar e comunicar os seus estudos [4]. A avaliação formativa pode ajudar na correção dos rumos, na adequação dos processos, podendo ser conduzida, pela contratação de especialista para comentar o plano da avaliação e ajudar na solução de problemas; ou na utilização de um painel orientador composto de avaliadores especialistas, membros da organização onde ocorre a avaliação. Essa opção pode ampliar a adesão da instituição às conclusões do processo avaliativo, uma vez que os envolve no processo. Pode-se ainda contar com observadores independentes atuando como meta-avaliadores, que revisam os relatórios parciais em andamento, buscando cobrir pontos deficientes [3,4]. As meta-avaliações somativas ocorrem após o término da avaliação, utilizando relatórios prontos, bancos de dados, documentação e registros relacionados a avaliação desenvolvida [4], e assim, auxiliará na percepção dos pontos fortes e as fragilidades da avaliação [6].

Ressalta-se que há distinção entre meta-avaliação e meta-análise. A meta-análise consiste em uma técnica estatística usada para combinar dados de múltiplos estudos, sobre uma mesma questão de pesquisa, combinando, em uma medida resumo, os resultados de tais estudos [3]. Enquanto a meta-avaliação, consiste no exame dos métodos e informações fornecidos por uma avaliação ou mais avaliações, visando o julgamento sobre a qualidade, mérito e valor da avaliação, seguindo os padrões de avaliação [10,17].

Na condução da meta-avaliação os protocolos da Joint Committee (2011) [10], indicam o emprego de listas de

verificação para facilitar a compreensão e o julgamento das diferentes avaliações. Recomenda-se empregar escalas padronizadas, que podem ser reproduzidas, buscando minimizar as diferenças atribuídas ao avaliador. Essas escalas podem conter critérios de julgamento, destacando o que foi atendido, parcialmente atendido, não atendido ou não se aplica. Assim, diferentes julgadores podem emitir seus valores, o que facilita a comparação, além de confrontar as discrepâncias [10]. Outra etapa consiste em discutir os resultados obtidos com a equipe do programa e os demais interessados, para se ter um consenso. A aplicação dos princípios e padrões avaliativos busca o maior envolvimento dos interessados no processo avaliativo, a explicitação dos conflitos de interesses, a ampliação da credibilidade do processo avaliativo, o suporte nas melhores evidências, a proteção de informações confidenciais e o uso de metodologias adequadas, que garantam a qualidade dos relatórios e conclusões [21].

A JCSEE define 30 padrões para meta-avaliações, entretanto o critério de seleção dos padrões pode variar de acordo com o que é considerado relevante para julgar a qualidade técnica das avaliações. A quantidade de padrões pode variar conforme os objetivos, do contexto e do escopo da avaliação, por exemplo, no estudo de metaavaliação do Plano Nacional de Desenvolvimento Sanitário (PNDS) da República da Guiné-Bissau foi utilizado 15 dos 30 padrões preconizados pela JCSEE [8] e na meta-avaliação do *after-school programs* utilizaram 17 padrões [17]. Assim, os padrões podem ser adaptados as questões específicas de cada processo avaliativo [6].

A aplicação dos princípios e padrões avaliativos busca o maior envolvimento dos interessados no processo avaliativo, a explicitação dos conflitos de interesses, a ampliação da credibilidade do processo avaliativo, o suporte nas melhores evidências, a proteção de informações confidenciais, e o uso de metodologias adequadas, que garantam a qualidade dos relatórios e conclusões [21]. Diferentes técnicas podem ser empregadas para se gerar o consenso, desde as conversas, explicitações dos pontos de vista, das divergências e convergências, a

soma dos escores e a definição de médias, até o emprego de grupos Delphi [33].

A meta-avaliação pode ser conduzida a partir da revisão de literatura e aplicados padrões, a partir do material identificado. Quando pertinente, o meta-avaliador pode ampliar as informações para a melhor definição dos escores. Assim, além dos relatórios existentes, pode-se também utilizar bases de dados coletados, documentação analisada, registros feitos e depoimentos obtidos e todo o material disponível que retrate a avaliação desenvolvida [34]. Cabe também, quando necessário, utilizar nova coleta primária mediante entrevistas com os atores chaves, usuários da avaliação e outros [8]. Ademais, a atribuição dos escores, deverá considerar todo o material e, quando necessário, utilizar informações quantitativas e qualitativas. Na meta-avaliação do Plano em Guiné Bissau, conduziram análise de conteúdo de Bardin, confrontando os relatórios disponíveis, com entrevistas coletadas pelos meta-avaliadores, para complementar o processo, o que aumenta a capacidade analítica do avaliador e sua capacidade de orientar os processos [8].

Conclusão

A meta-avaliação pode ser uma ferramenta útil, em especial, quando se trata de temas e contextos que já avançaram na produção de avaliações prévias. A meta-avaliação contribuirá no julgamento da qualidade dos processos avaliativos, na adequação e utilidade da sua aplicação. Deve ser conduzida por avaliadores experientes, adaptado ao contexto local. Assim, a aplicação dos padrões não deve ser mecânica, mas adaptada por cada avaliador ao seu contexto. Destaca-se ainda a sua articulação com a gestão, estabelecendo um julgamento de valor, visando subsidiar eventuais tomadas de decisão. Essencialmente o que se busca é definir a utilidade do programa, do objeto avaliado e como poderá apoiar o processo de decisão.

Bibliografia

1. Hartz ZMA, Contandriopoulos AP. Integralidade da atenção e integração de serviços de saúde: desafios para avaliar a implantação de um “sistema sem muros”. *Cad. Saúde Pública*. 2004; 20(Suppl 2): S331-S336.
2. Machado TR, Chaise RM, Elliot LG. A meta-avaliação como instrumento de qualidade nas Políticas Públicas: o Programa Segundo Tempo. *Rev Meta Aval*, 2016; 8: 1-20.
3. Davidson EJ. *Evaluation methodology basics*. Thousands Oaks, CA: Sage, 2005.
4. Elliot LG. Metaevaluation: from approaches to possibilities of application. *Ensaio: Avaliação e Políticas Públicas em Educação*, 2011; 19(73): 941-64,.
5. Hartz Z. Meta-avaliação da atenção básica à saúde – teoria e prática. Rio de Janeiro: FIOCRUZ, 2008, 410p.
6. Stufflebeam DL. The metaevaluation imperative. *American Journal of Evaluation*, 2001; 22(2): 183-209.
7. Hartz ZMA. Principles and standards in metaevaluation: guidelines for health programs. *Ciência & Saúde Coletiva*, 2006; 11(3): 733-38.
8. Guerreiro CS, Ferrinho P, Hartz ZMA. Avaliação em saúde na República da Guiné-Bissau: uma meta-avaliação do Plano Nacional de Desenvolvimento Sanitário. *Saúde em Debate*, 2018; 42: 549-65.
9. Silveira R. Um exame da organização textual de ensaios científicos. *Revista Estudos Linguísticos*, 1992; 21: 1244-52.
10. Joint Committee on Standards for Educational Evaluation. *The program evaluation standards: a guide for evaluators and evaluation users: a guide for evaluators and evaluation users*. 3rd. ed. Thousands Oaks, CA: Sage, 2011.
11. Fitzpatrick JL, Sanders JR., Worthen BR. *Program Evaluation: Alternative Approaches and Practical Guidelines*, 4th ed. Boston: Pearson Education, 2011.
12. Cooksy LJ, Caracelli VJ. Meta-evaluation in Practice. *Journal of multidisciplinary evaluation*, 2009; 6(11): 1-15.
13. Hartz ZMA. Meta-avaliação da gestão em saúde: desafios para uma “nova saúde pública”. *Ciência & Saúde Coletiva*, 2012; 17: 832-34.
14. Scriven M. *Evaluation thesaurus*. 4th. ed. Newbury Park, CA: Sage, 1991.
15. Khalil H. et al. An evidence – based approach to scoping reviews. *Worldviews on Evidence – Based Nursing*, 2016; 13(2): 118-23.
16. Figueiró AC, Hartz ZMA, Thuler LCS, Dias ALF. Meta-evaluation of baseline studies of the Brazilian Family Health Strategy Expansion Project: a participatory and formative approach. *Cad. Saúde Pública*. 2010; 26(12): 2270-78.
17. Scott-Little C, Hamann MS, Jurs SG. Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation*, 2020; 23(4): 387-419.
18. Joint Committee on Standards for Educational Evaluation. *The program evaluation standards: how to assess evaluation of educational programs*. Hightstown, NJ: McGraw-Hill Book Co., 1981.
19. Joint Committee on Standards for Educational Evaluation. *The program evaluation standards: how to assess evaluation of educational programs*. 2nd. ed. Thousands Oaks, CA: Sage, 1994.
20. Joint Committee on Standards for Educational Evaluation. *Checklist of The Program Evaluation Standards Statements*. Western Michigan University, 2018.
21. Hartz ZMA et al. Multi-strategy in the Evaluation of Health Promotion Community Interventions: An Indicator of Quality. In: Potvin L, McQueen D. (eds.), *Health Promotion Evaluation Practices in the Americas*, DOI: 10.1007/978-0-387-79733-5 14, C Springer Science Business Media, LLC, 2008.
22. Bardin L. *Análise de conteúdo*. São Paulo: Edições 70. 2011.
23. Scriven M. *Evaluation thesaurus*. 4th. ed. Newbury Park, CA: Sage, 1991.
24. Stufflebeam DL. *Program Evaluations Summary Metaevaluation Checklist (Based on The Program Evaluation Standards)*. 2012.
25. Fronteira I, Ferrinho F, Dussault G, et al. Relatório de Avaliação Final da Execução do Plano Nacional de Desenvolvimento Sanitário 2003-2007 da República da Guiné-Bissau. Relatório de Avaliação Final, Associação para o Desenvolvimento e Cooperação Garcia de Orta, 2007.
26. Brasil, Ministério da Saúde. *Pública da República da Guiné Bissau. Plano Nacional de Desenvolvimento Sanitário 1998-2002*. 1997.
27. Brasil, Ministério da Saúde Pública da República da Guiné-Bissau. *Plano Nacional de Desenvolvimento Sanitário II 2008-2017*. 2007.
28. Hartz ZMA. Meta-evaluation of health management: challenges for ‘new public health’. *Ciênc Amp Saúde Coletiva* 2012; 17: 832–834.
29. Isabella Samico, Eronildo Felisberto, Ana Claudia Figueiró, et al. Avaliação em Saúde: bases conceituais e operacionais. Medbook, Editora Científica. Rio de Janeiro, 2010.
30. Furtado JP, Laperrière H. Parâmetros e paradigmas em meta-avaliação: uma revisão exploratória e reflexiva. *Ciência & Saúde Coletiva*, 2012; 17: 695-705.
31. Cotera AT, Matamoros AA. Propuesta para implementar la Metaevaluación en la Escuela de Ciencias de la Educación de la Universidad Estatal. *Revista Calidad en la Educación Superior*, 2011; 2(1): 247-65.
32. Worthen BR, Sanders JR, Fitzpatrick JL. *Avaliação de programas: concepções e práticas*. São Paulo: Ed. Gente, 2004.
33. Vieira LMS, Souza LEPF, Hartz ZMA. Conferência de consenso sobre a imagem-objetivo da descentralização da atenção à saúde no Brasil. In: Vieira LMS, Hartz ZMA. (Org.). *Avaliação em saúde: dos modelos teóricos à prática na avaliação de programas e sistemas de saúde*. 1ed. Rio de Janeiro; Salvador: FIOCRUZ/EDUFBA, 2005, v. , p. 65-102.
34. Dias Sobrinho J. *Avaliação: políticas educacionais e reformas da educação superior*. São Paulo: Cortez, 2003.

Conflitos de interesses:

Os autores declaram que não existem conflitos de interesses.