

Health evaluation in the Republic of Guinea-Bissau: a meta-evaluation of the National Health Development Plan

Avaliação em saúde na República da Guiné-Bissau: uma meta-avaliação do Plano Nacional de Desenvolvimento Sanitário

Cátia Sá Guerreiro¹, Paulo Ferrinho², Zulmira Hartz³

DOI: 10.1590/0103-1104201811801

ABSTRACT In the Republic of Guinea-Bissau, one of the poorest countries in the world, the National Health Development Plan (NHDP) is intended to be the national health guidance document. NHDP II (2008-2017) succeeded NHDP I (1998-2002, revised for 2003-2007), being both assessed. In 2017, a new planning process – NHDP III – was completed. As for NHDP, a meta-evaluation was carried out, aiming to assess the quality of the evaluations carried out and to verify the use of evaluations' results in subsequent interventions. Applying meta-evaluation patterns to evaluation reports, it was verified that, as to the four principles – utility, feasibility, adequacy and accuracy –, both evaluations presented a satisfactory score higher than 66.6%, showing an excellence character in various patterns. Methodologically, NHDP I report stands out from NHDP II. In the planning phase, the suggestions made by the previous evaluation were reconsidered for both NHDP II and III. The fragility of the Country emerges as a pattern of specificity, revealing that NHDP has offer a reference in scenarios of instability. These evaluations fit the current challenge of considering the context and the culture of peoples as an integral part of the evaluation.

KEYWORDS Evaluation studies as subject. Implementation of health plan. Planning. Culture.

RESUMO Na República da Guiné-Bissau, um dos países mais pobres do mundo, o Plano Nacional de Desenvolvimento Sanitário (PNDS) pretende ser o documento de orientação nacional em saúde. O PNDS II (2008-2017) sucedeu ao PNDS I (1998-2002, revisto para 2003-2007). Ambos foram alvo de avaliações. Em 2017, completou-se um novo processo de planejamento, o PNDS III. Considerando o PNDS, procedeu-se a uma meta-avaliação objetivando avaliar a qualidade das avaliações efetuadas e verificar a utilização dos resultados das avaliações em intervenções subsequentes. Aplicados padrões de meta-avaliação aos relatórios de avaliação, verificou-se que, relativamente aos quatro princípios de meta-avaliação – utilidade, exequibilidade ou factibilidade, propriedade e precisão ou acurácia –, ambas as avaliações apresentam a classificação de 'satisfatório' (score superior a 66,6%), revelando um caráter de excelência para vários padrões. Metodologicamente, o relatório referente ao PNDS I sobressai sobre o do PNDS II. Na fase de planejamento, as sugestões deixadas pela avaliação anterior foram levadas em conta tanto para o PNDS II como para o III. A fragilidade do Estado da Guiné-Bissau emerge como padrão de

¹Universidade Nova de Lisboa (UNL), Instituto de Higiene e Medicina Tropical (IHMT), Global Health and Tropical Medicine (GHTM) - Lisboa, Portugal.
Orcid: <https://orcid.org/0000-0002-0505-2155>
cs.guerreiro22@gmail.com

²Universidade Nova de Lisboa (UNL), Instituto de Higiene e Medicina Tropical (IHMT), Global Health and Tropical Medicine (GHTM) - Lisboa, Portugal.
Orcid: <https://orcid.org/0000-0002-3722-0803>
pferrinho@ihmt.unl.pt

³Universidade Nova de Lisboa (UNL), Instituto de Higiene e Medicina Tropical (IHMT), Global Health and Tropical Medicine (GHTM) - Lisboa, Portugal.
Orcid: <https://orcid.org/0000-0001-9780-9428>
zhartz@ihmt.unl.pt



especificidade. O PNDS tem sido uma referência em cenários de instabilidade. Tais avaliações foram ao encontro do atual desafio de considerar o contexto, e, nele, a cultura dos povos, como parte integrante da avaliação.

PALAVRAS-CHAVE *Estudos de avaliação como assunto. Implementação de plano de saúde. Planejamento. Cultura.*

Introduction

Republic of Guinea-Bissau, a fragile State

Located westerly of the African continent and comprising a continental belt and the Bijagós Archipelago, the Republic of Guinea-Bissau (RGB) is an ancient Portuguese colony that became independent in 1973, after 15 years of armed struggle.

Since the proclamation of independence, the country has experienced situations of political and institutional instability, materialized by repeated coups and armed conflicts. This scenario has involved frequent changes in the leadership of teams responsible for the various ministries. Two situations of particular instability in the country's history are to be highlighted: the political-military conflict occurred between June 1998 and May 1999, that particularly impacted on the destruction of infrastructure and social¹ system, and the coup d'état taken place in 2012, specially impacting on the political stability and the national economy until today.

According to the '2015 Index of Fragile States', published by The Fund for Peace², RGB is considered the 17th most fragile country in the world among 178 countries, where government limitations and provision of public goods and services to citizens fall into the

criteria of lack of 'legitimacy' and 'effectiveness' defined by the Organization for Economic Co-operation and Development (OECD) for the Characterization of Fragile State (FS)³.

FS Health Systems (HS), due to their various needing scenarios, can be 'flooded with help' in a variety of formats, including international Non-Governmental Organizations (NGOs) and bilateral and multilateral agencies, often collaborating with local civil society organizations, although they are not always able to understand the implications of their interventions in the specific context of FS⁴. Alongside the public-private partnerships involved in the health sector, there is also a large number of global health initiatives created to be one of the globalization benefits, which are programs usually geared towards specific diseases⁵. Local officials are tasked with managing the assistance, not always in a planned way, provided by these multiple stakeholders.

As described above, RGB presents a context of fragility and shortages in the health sector and its HS operation. In the last two decades, the percentage of the State General Budget (OGE) allocated to Health was less than 10%, which represents less than 15% of the sector's financial needs, implying the dependence on external support⁶. As for the World Health Organization (WHO)⁷, about 90% of the sector funding have been ensured by cooperation partners.

The Strategic Health Planning in RGB

RGB has been an independent state since 1973. It established its National Health Policy (NHP) in 1993 and has practiced Strategic Health Planning (SHP) over the years.

In order to define a framework for sanitary development activities and actions based on a sector development policy, the first major SHP exercise in RGB was carried out in collaboration with developing partners, resulting in the National Health Development Plan I (NHDP I), stated for the period 1998-2002⁸. However, the 1998 political-military situation undertaken the plan implementation, which was revised for the period 2003-2007 to allow for the implementation of the activities initially planned^{6,9}.

In response to the challenges posed by the National Strategy Document on Poverty Reduction II (Denarp II)¹⁰, and to carry on the path gone through NHDP I implementation, NHDP II was created for the period 2008-2017⁶. Since its ordeal in 2008, NHDP II has remained as the reference framework for the health development activities and actions, although its implementation has been compromised by the political-military situation experienced by RGB¹¹. In March 2017, a new planning process took place, resulting in the NHDP III drafting, due to the decision of not extending NHDP II implementation period. The third NHDP is oriented on the NHP adopted in March 2017, which replaces that of 1993 and still awaits for validation by the Council of Ministers, scheduled for early 2018 by the Strategic and Operational Plan of Terra Ranka Government 2015-2025¹²; and by the recommendations emerging from the I National Health Conference in October 2014.

Evaluation has been an integral part of the SHP process in RGB. There is already a broad consensus around the idea that public policies should be accompanied by systematic evaluation as part of government routine, continuously adapting to their results and

recommendations. This idea implies the need to evaluating, but it is necessary to continuously question the evaluation ability to produce the information and judgment necessary to help improve HS performance¹³.

Objectives

Having the NHDP^{6,8} as the national strategic orientation documents for the health sector, and since evaluation data on its implementation are approved and available^{9,11}, a meta-evaluation was carried out with a twofold objective: to evaluate the quality of evaluations carried out and to know to what extent evaluations' results were or are being used in further interventions. Therefore, the aim is to contribute to the increased credibility of the evaluation process within SHP scope in RGB, contextualizing its role and enhancing its usefulness for decision making in the health sector.

Material and methods

The meta-evaluation was carried out by means of content analysis¹⁴ regarding NHDP I and II implementation evaluation documents, and eleven interviews¹⁵ conducted with SHP evaluators and key stakeholders in RGB.

The eleven interviewees were selected by the investigator for holding intended information for the study, consisting, therefore, of a non-probabilistic intentional sampling¹⁶. Two of them are evaluators, one is involved in NHDP I evaluation and the other in both evaluation processes, as well as in NHDP II and III drafting. The nine key actors, all Guineans, directly participate or participated in the health planning processes since NHDP I elaboration till NHDP III current drafting. They hold or have held government positions in the sector at the national level, two of them being

strongly related to specific programs, i.e. Fight against Tuberculosis and Leprosy and Maternal and Child Health. Two respondents currently hold positions in international organizations. The interviews were recorded with the interviewees' agreement.

Meta-evaluation can be plainly defined as the evaluation of the evaluation, and, operationally speaking, as a process of description, judgment and synthesis of studies or evaluation procedures aiming at assuring the quality of the studies on evaluation. By providing a systematic analysis of an evaluation study¹⁹, its methodology allows to verify if the objectives initially set in the evaluation were adequately and effectively achieved, and whether the development of the program, project or product revealed its merit²⁰. It thus leads to a practice of thinking over all the procedures used in the evaluation, creating opportunities for the incorporation of new knowledge by not only meta-evaluators but also by evaluators¹⁸, and providing aid to decision-makers as to improve the performance of their actions¹⁷.

Thus, meta-evaluation allows access to information about the limitations and potential of the evaluation carried out, increasing its credibility and allowing stakeholders to judge and contextualize the results obtained²¹.

Appropriate procedures for conducting a meta-evaluation may vary according to the type of evaluation performed²². The literature suggests some possibilities for conducting a meta-evaluation, and the most well-known conceptual framework is that produced by the Joint Committee on Standards for Educational Evaluation

(JCSEE) in 1981. It was updated in 1994, and its 2011 version is actually in force^{18,22,23}. A total of 30 evaluation patterns are organized around four principles: 'Utility' – meets stakeholder information needs; 'Feasibility' – is realistic and moderate in resources and costs so to justify its implementation; 'Adequacy' – is conducted ethically, with respect for stakeholders; and 'Precision and accuracy' – discloses and transmits information about the value or merit of the programs assessed within due date^{18,24}.

To ease the task of judging the quality of evaluations, JCSEE built a checklist of patterns broken down into categories to which a scale can be applied, allowing to judge the fit level of the evaluated object in relation to each pattern^{17,22}. The literature adopted ensures that, according to the nature of the evaluation or meta-evaluation, the application of patterns varies, because, depending on the object of interest, some of the parameters may not be adequate or applicable²².

Among the 30 JCSEE meta-evaluation patterns, this study chose thirteen to use in the analysis of evaluation reports for the (i) Principle of Utility – Assessor credibility, Clearness of reports, Evaluation impact; (ii) Principle of Feasibility – Practical procedures, Contextual feasibility; (iv) Principle of Adequacy – Fair and complete evaluation, Dissemination of results, Conflict of interest; and (v) Principle of Precision or Accuracy – Content analysis, Description of objectives and procedures, Reliable sources of information, Justifiable conclusions, Impartiality of reports. *Chart 1* describes the textual definition of the patterns applied, drawn up on the basis of the literature adopted^{17,23-25}.

Chart 1. Textual definition of the principles and patterns applied

PRINCIPLE	PATTERN	TEXTUAL DEFINITION
UTILITY Meets the stakeholder information needs.	Evaluator credibility	Evaluations should be conducted by qualified persons who establish and maintain the evaluation credibility.
	Clearness of reports	Evaluation reports should clearly describe the assessed program, including its context and the objectives, procedures and conclusions of the evaluation, so as to provide essential information that is easily understood.
	Evaluation impact	Evaluation should be planned, conducted and disseminated in a way that encourages stakeholder monitoring, enhancing the possibility of using evaluation results.
FEASIBILITY The evaluation is realistic, conducted without unnecessary spending, boosting the assessment effectiveness and efficiency.	Practical procedures	Evaluation procedures should be practical so as to avoid disturbance when collecting information.
	Contextual feasibility	Evaluations should recognize, monitor and balance interests and cultural and political needs of individuals and groups. The evaluation should be planned and conducted so as to anticipate different positions of different stakeholders, achieving the cooperation of all.
ADEQUACY The evaluation is conducted ethically, with respect by those involved.	Fair and complete Evaluation	Evaluations should provide full descriptions of main positive points, limitations and conclusions of the assessed program, allowing the recovery of success aspects and the correction of existing failures.
	Dissemination of results	Those responsible for the Evaluation shall ensure that all stakeholders subject to the evaluation process or legally entitled to it have access to the evaluation results in its entirety, unless such access violates legal and proprietary obligations.
	Conflict of interest	Evaluations should openly and honestly identify and address actual or perceived conflicts of interest that might harm the evaluation.
PRECISION AND ACCURACY Disclose and transmit information about the value or merit of programs assessed within due date.	Content analysis	The context in which the program is inserted should be described with the necessary detail so that its possible influences on the program can be identified.
	Description of objectives and procedures	The objectives and procedures of evaluation should be clearly described and monitored so that they can be easily identified and examined.
	Reliable sources of information	The sources of information used in the evaluation should be described in detail so as to allow for the adequacy analysis of the information collected.
	Justifiable conclusions	The conclusions of an evaluation should be explicitly justified so that they can be analyzed by the key stakeholders or affected by the evaluation or the program.
	Impartiality of reports	Reporting procedures should include methods to prevent possible distortions caused by personal feelings or biases so that the reports fairly reflect the results of the evaluation.

A 0-10 quantitative scale was applied to each of the patterns, whereby the values 9 and 10 correspond to 'excellent'; 7 and 8, to 'very good'; 5 and 6, to 'good'; 3 and 4, to 'weak; and 0-2, to 'critical'. A re-qualification by principle of evaluation was further carried out, split in 'Unsatisfactory', 'Acceptable' and 'Satisfactory', following layers <33.3%; 33.3%-66.6% and >66.6%, respectively^{13,17,19}.

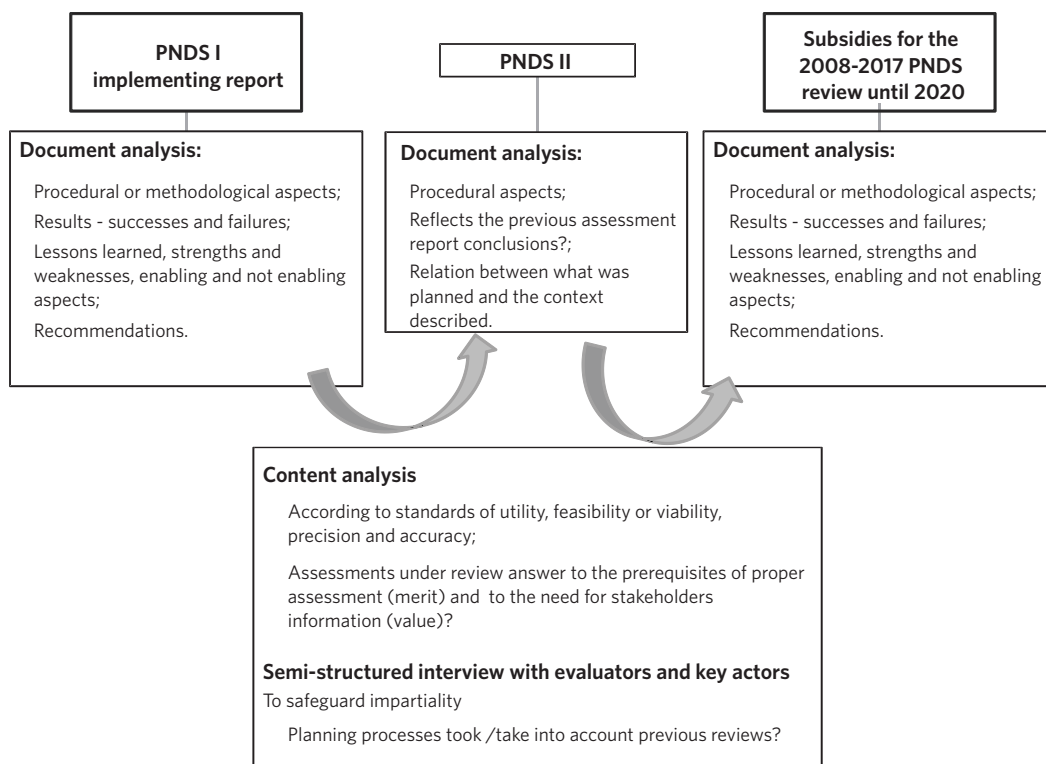
As for the ranking by pattern and, thus, the calculation of value by principle of meta-evaluation, content analysis of available evaluation reports was carried out.

It was also sought to understand the extent to which the results of evaluations carried out were or are being used in subsequent interventions, further proceeding to the content analysis of the planning

document that followed the evaluation – the NHDP II. The information obtained was complemented by means of content analysis of interviews with evaluators and other key actors. Because it is a circumstantial dynamic process, NHDP III drafting create a new objective to interviews that was to

obtain subsidies on the planning process, specifically on the utility and impact of previous evaluations. Because one of the evaluators was or is involved in planning, one of the interviews was also intended to safeguard impartiality. The methodology applied is summarized in *figure 1*.

Figure 1. Methodological map



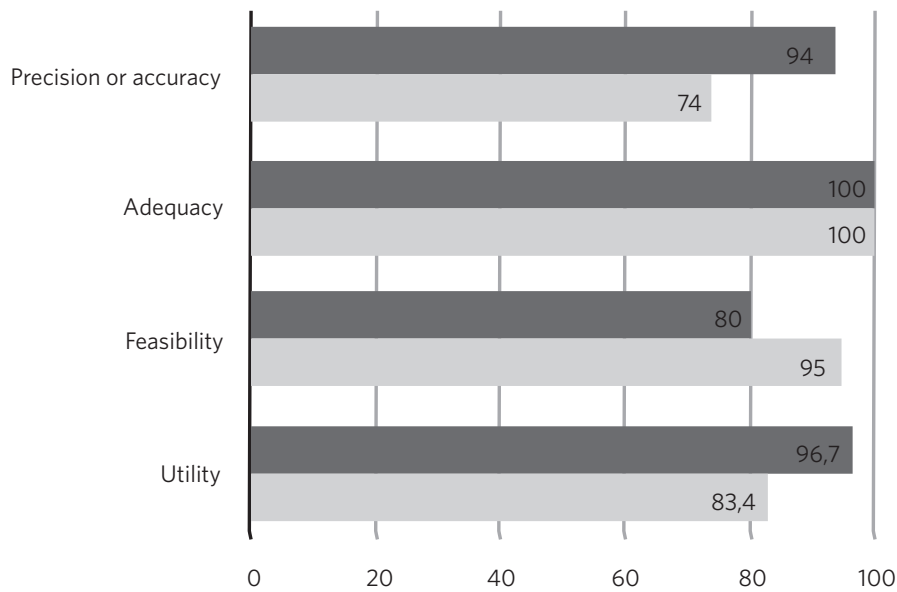
Results and discussion

Comparing evaluation reports

Based on the meta-evaluation carried out

by patterns, it is possible to verify that, as to the four principles – Utility, Feasibility, Adequacy, and Precision or Accuracy –, both evaluations are classified as ‘satisfactory’, because they scored above 66.6% (*figure 2*).

Figure 2. Application of meta-evaluation principles (%)



Concerning each evaluated pattern (figure 3), both reports present a character of excellence in several patterns, that is, ‘assessor credibility’, ‘evaluation impact’, ‘contextual feasibility’, ‘fair and complete

evaluation’, ‘dissemination of results’, ‘conflict of interests’, ‘content analysis’, and ‘justifiable conclusions’. Also for the patterns alluding to the ‘Principle of Adequacy’, both reports present a character of excellence.

Figure 3. Assessment by meta-evaluation standard

		R1	R2
Utility	Evaluator credibility	Excellent	Excellent
	Clearness of reports	Excellent	Very good
	Evaluation impact	Excellent	Excellent
Feasibility	Practical procedures	Very good	Excellent
	Contextual feasibility	Excellent	Excellent
Adequacy	Fair and complete assessment	Excellent	Excellent
	Dissemination of results	Excellent	Excellent
	Conflict of interest	Excellent	Excellent
Precision or Accuracy	Content analysis	Excellent	Excellent
	Description of objectives and procedures	Excellent	Very good
	Reliable sources of information	Excellent	Very good
	Justifiable conclusions	Excellent	Excellent
	Impartiality of reports	Very good	Very good

Excellent
Very good
Good
Weak
Critical

Charts 2, 3, 4 and 5 describe the rationale behind the classification of the patterns.

Chart 2. Principle of utility, evaluation by patterns

Principle of Utility	PNDS I Evaluation Report		PNDS II Evaluation Report		Note
	Pattern	Ranking value Justification	Ranking value Justification		
Ranking	R1=29 = 96,7%	Final ranking: Satisfactory	R2=25 = 83,4%	Final ranking: Satisfactory	
Evaluator's credibility	10	The team coordination of evaluators had an extensive knowledge of the contextual reality and of the PNDS I deployment process.	10	The evaluator followed the process of drafting and implementing PNDS II, and carried an extensive knowledge of the contextual reality. <i>"The fact that I know PNDS II and the context of its implementation is an advantage"</i> , says the evaluator.	
Clearness of reports	10	The report clearly describes the program assessed, including the evaluation context, objectives, procedures and conclusions.	5	Very long report, failure to systematize information, poor referencing. Very long introduction. Clearly describes objectives and context. Very explicit conclusions, but without clearly defined procedures.	Information obtained solely through document analysis of evaluation reports.
Evaluation impact	9	The evaluation was followed up by local stakeholders, members of Minsap. The subsequent PNDS II planning process <i>"was based not only on the evaluation report of PNDS I implementation but also on an external consultancy from WHO-Geneva [...]".</i> The team coordination of PNDS I evaluators <i>"was consulted throughout the process [...]".</i>	10	Although taking place in a troubled period of RGB history, since <i>"during the assessment period, three different health ministers occupied the position and none of them was able to make decisions"</i> , the whole process <i>"was strongly supported by the national director of public health"</i> . The evaluation process enabled <i>"the dialogue between partners and Minsap leaders"</i> as well as <i>"the perception of adjustments to be made so that it could finally be deployed"</i> - it was a document that supported the further option of not reformulating PNDS II, but to draft a new plan - the PNDS III.	Information obtained through the interviews.

Chart 3. Principle of feasibility, evaluation by patterns

Principle of Feasibility	PNDS I Evaluation Report		PNDS II Evaluation Report			
	Pattern	Ranking value	Justification	Ranking value	Justification	Note
Ranking	R1=16 = 80%	Final ranking: Satisfactory	R2=19 = 95%	Final ranking: Satisfactory		
Practical procedures	7	Although very clearly defined in the evaluation report, and thus well-defined at an evaluation preparation stage, the procedures proved to be difficult to implement at the time of field evaluation. They were prepared in Lisbon, based on the PNDS I document, from which it was intended to assess the implementation. However, a civil war occurred and "upon reaching the ground, the data able to calculate the indicators had disappeared". The procedures proved not to be as practical as they seemed because they were misfit of local reality at the time of the evaluation.	9	Although the procedures were not clearly described in the report based on the interview with the evaluator, it was possible to verify that they responded to the evaluation needs and fit the difficult context then experienced in the country and in the health sector. "The evaluation was conducted in a very difficult context [...]. Due to the disorganization that existed, I took the initiative to call and schedule the meetings [...]. I sought to confront the information obtained in meetings and interviews with the quantitative data I was able to access". From the interview, one can conclude that procedures were practical and allowed to retrieve data contained in the genesis of the assessing document, although the recording of that procedural approach is flawed.	Additional information obtained through the interviews.	
Contextual feasibility	9	"The request for evaluation comes from Minsap and not from any of the donors or partners", not excluding the existence of different stakeholders in the program being assessed. Evaluation process was supported and motivated by Minsap.	10	"The aim was to try to hear as many national leaders as possible, not only those from Minsap but also the institutional, programmatic and regional ones. [...] I also heard funders, because their perspective is very important. It was still possible to conduct a discussion process; plenary sessions took place between health leaders and industry partners [...]".	Information obtained through the interviews.	

Chart 4. Principle of adequacy, evaluation by patterns

Princípio de Propriedade	Relatório Avaliação PNDS I		Relatório Avaliação PNDS II			
	Pattern	Ranking value	Justification	Ranking value	Justification	Note
Ranking	R1=30 = 100%	Final ranking: Satisfactory	R2=30 = 100%	Final ranking: Satisfactory		
Fair and complete assessment	10	Chapters 4 and 5 of the report focus exclusively on the description of the strengths, limitations and conclusions of the assessed program, leaving some suggestions and recommendations to Chapter 5.	10	In its chapters 3 and 4, the evaluation report outlines clearly the strengths, limitations and conclusions of the assessed program, leaving recommendations for the future.	Information obtained solely through document analysis of assessment reports.	
Dissemination of results	10	Those who requested the evaluation and participated in it were recognized in the report; the full disclosure of results to them was validated through the interviews..	10	Those who requested the evaluation and participated in it were recognized in the report; the full disclosure of results to them was validated through the interviews.	Information fully obtained through the interviews.	

Chart 4. (cont.)

Conflict of interest	10	In the executive summary of the assessed document, the team of evaluators is described, and the lack of conflicts of interest is added. The information was validated through the interviews together with the evaluators.	10	The close relationship between the evaluator and the program under evaluation could raise issues as of conflict of interest. However, in the executive summary, the evaluation objectives and the role of the evaluator are clear. The evaluator expresses, during an interview: <i>"The fact that I knew the reality and was involved helped me to bridge the gap between the reality experienced and the objective of the evaluation of which I was the protagonist, just that"</i> .	Additional information obtained through the interviews.
----------------------	----	--	----	---	---

Chart 5. Principle of precision and accuracy, evaluation by patterns

Principle of Precision	PNDS I Evaluation Report		PNDS II Evaluation Report			
	Pattern	Ranking value	Justification	Ranking value	Justification	Note
Ranking	R1=47 = 94%	Final ranking: Satisfactory	R2=37 = 74%	Final ranking: Satisfactory		
Content analysis	10	In chapter 1 of the report, the context is well defined and analyzed, and its influence can be retrieved from the results presented in the same chapter.	10	The context is extensively defined in chapter 1, with particular emphasis on health strategic planning.	Information obtained solely through document analysis of evaluation reports.	
Description of objectives and procedures	10	The predicted objectives and assessing procedures are clearly described in Chapter 2, as are described the way in which they were deployed.	5	In the executive summary and chapter 2 - objectives and methods, the objectives are clearly defined, while the procedures are not. Mention is made to data sources, but there is no detail in the description of the information retrieving procedures, preventing from examining the paths or recuperating them.	Information obtained solely through document analysis of evaluation reports.	
Reliable sources of information	10	Data sources are clearly defined, and considered appropriate because they are also involved in the PNDS implementation or in relevant documents related to the plan under analysis	7	Although the sources of information are identified and appropriate to the assessment process in question, given the context of political instability experienced, <i>"it was not possible to obtain data from all the stakeholders provided for"</i> . It is not possible to monitor the access to the sources initially planned due to lack of information in the report (see procedures, previous pattern)	Additional information as of report II obtained through the interviews.	
Justifiable conclusions	10	The conclusions were framed in the context and allow for a prospective analysis of the results. <i>"We tried to go beyond the indicators, we tried to relate to the MDGs [Millennium Development Goals]"</i> .	10	The conclusions were framed in the context and allow for a prospective analysis of the results. The work recommendations can be used in future health planning processes.	Additional information obtained through the interviews.	

Chart 5. (cont.)

Impartiality of reports	7	Although the interview gave the notion that the evaluator tended to be impartial in the evaluation performed, no strategy or method is described in the evaluation report to prevent distortions caused by personal feelings or biases, and the detailed description of the evaluation procedures contributes to impartiality. <i>"We intended to convey what had to be said and do it in such a way as to create a lever, recognizing what had not been done and what needed to be done",</i> says one of the evaluators.	5	Although the interview gave the notion that the evaluator tended to be impartial in the evaluation performed, no strategy or method is described in the evaluation report to prevent distortions caused by personal feelings or biases. This fact is aggravated by the lack of clearness in the definition of the evaluation procedures. <i>"I do not agree with some of the recommendations that I give. They result from and reflect what was said during the interviews",</i> stresses the interviewee, safeguarding impartiality.	Analysis based mainly on the reports, mitigated by data obtained in an interview with the evaluators.
-------------------------	---	--	---	---	---

The excellence of 'context analysis' in both evaluation reports is revealing of the close relationship between evaluators and the object under study. However, due to procedural and methodological aspects, NHDP I report stands out on NHDP II'. In both evaluations, a very well-contextualized analysis of the reality lived in the RGB is deployed, but the interim report on subsidies for the revision of NHDP II, although defining its objectives, does not clearly describe the procedures of the evaluation process, not allowing, for example, its replication by another team of evaluators. Assuming that the assessor used reliable sources of information, these are neither identifiable in the report nor described in detail.

Reporting impartially is possible if reporting procedures include methods to prevent possible distortions caused by personal feelings, opinions or bias¹⁸. Above all, the interim report on subsidies for the revision of NHDP II does not describe the methods that prevent distortion, although they had arisen in the first evaluation, when evaluation procedures were described in more detail.

It turns out, however, that the aim of the second evaluation process was not to carry out the final evaluation of a plan or program implementation, but to gather subsidies for the decision-making on a NHDP' reformulation or drafting for RGB. Thus, the impact of

such failures is mitigated by the very reason of the evaluation.

As for the pattern 'practical procedures', it is concluded that, in relation to NHDP I implementation evaluation, these procedures, though well described and feasible, show some implementation risk if we consider the context experienced by RGB on the evaluation date. According to the evaluators, to assess NHDP I implementation,

The procedures were defined in the stage prior to the move to RGB, being done the complete planning of the evaluation process and its procedures, included, for example, the schedule of the information collection.

The work of evaluation preparation was done without full knowledge of RGB reality, which differed from the one at the time NHDP was written. After all, between NHDP I drafting and its implementation evaluation, a civil war took place with far-reaching consequences for the health sector. According to one of the evaluators interviewed, at the time of the evaluation, it happened that:

We carried out an evaluation structure, but, upon the arrival on the spot, we noticed that data to calculate the indicators had disappeared. So those goals ceased to make sense. There was a setback; everything was missing. We lacked

HR, pharmacological resources, infrastructure. We had to adapt the evaluation structure to that new context.

As for the same pattern ‘practical procedures’, the second report is briefly described in the body of the text. The interview assessor explains the way he implemented the evaluation, that is, adapting the procedures to the reality lived at the time of the interview:

The evaluation was carried out in a very difficult context. At the time of the evaluation, three health ministers held the position, though none of them was able to make decisions. The evaluation was accompanied by one of the national directors, greatly supported by the national director of public health. Due to the disorganization, I even took the initiative to call and set up meetings. The aim was to try to hear as many national leaders as possible, not only those coming from Minsap but also institutional, programmatic and regional ones. I also heard the funders, because their perspective is very important. [...]. It was also possible to conduct a discussion process; there were plenary sessions between health leaders and industry partners [...]. I sought to confront the information obtained from the meetings and interviews with the quantitative data I could access.

Considering the pattern ‘clearness of reports’ – according to which reports should clearly describe the assessed program, including its context and objectives, procedures and evaluation conclusions, so to provide essential information that is easily understood¹⁸ – NHDP I evaluation and implementation report is clearer than the interim report on subsidies for the revision of NHDP II. Having safeguarded the asymmetry concerning the objective of the evaluation process, it is important to emphasize that both reports show excellence as regards to the justification of the conclusions so that

they can be analyzed and used later, conferring to the pattern ‘justifiable conclusions’ a ranking of excellence for both evaluations.

Finally, with respect to the pattern ‘fair and complete evaluation’, it is verified that both the positive and the negative aspects are listed in both reports as well as suggestions based on the in-depth analysis of the context, realistically approaching each of the NHDP’ strengths and weaknesses.

Illustrating what has been described, it can be read in NHDP I evaluation report:

It can thus be concluded that health services [...] continue to fail to ensure care and quality [...] although in some cases a slight improvement and a measurable effort are made to provide the country with health infrastructure. It is necessary, however, to think about the future, the reason why we suggest [...]. Intersectoral collaboration and partnership building have fallen short of what was expected, being necessary that [...].

In the same sense, it can be read in the interim report on subsidies for the revision of NHDP II:

In a simplistic way, there has been significant progress in two axes of NHDP [...], some poorly sustained progress in four axes [...] and very unsatisfactory results in two axes [...]. Despite the political instability that worsened its implementation, NHDP II had the merit of remaining as the guiding document of the health sector and of contributing to some stability of an essential sector for the development of Guinean society. [...] As final recommendations, we prioritize the following [...].

As for the second objective of this meta-evaluation, such evaluations were largely taken into account in the planning process that followed. Regarding the use of their data in the planning stage, the suggestions made by the previous evaluation were also taken into account in the preparation of both NHDP II and NHDP III. At both instances,

evaluation reports were key documents for the planning process. *“The results of NHDP implementation evaluation were and are widely considered in health planning processes”*, says one of the key stakeholders interviewed. Another key actor, referring to the process of drafting NHDP II, states that

the previous evaluation was taken into account; the committees were established on the basis of the recommendations contained in the evaluation document.

Addressing the SHP process that resulted in the NHDP III drafting, one interviewee explains: *“As to the evaluation report on NHDP II implementation, decision was made to draw up a new plan, to start over”*.

Thinking over the results

The relation between evaluators, assessed plans and SHP process in RGB deserves open interest and attention also by verifying impartiality as by ensuring the credibility of the meta-evaluation, based on the fact that each evaluation study is likely to be biased. The decisions an assessor makes about what he examines – methods, instruments, groups to hear etc. – impacts the evaluation outcome²¹. The same authors affirm that the assessor’s own personal history, preferences, and experience affect the way to conduct the study. In this meta-evaluation, this reality became evident, although, as far as possible, impartiality was safeguarded, ensuring credibility by means of the methodology that followed the literature adopted. The fact that this meta-evaluation took into account not only the content of reports but also the information obtained from the evaluators directly involved, from others displaying some externality, from evaluation users, and from key players in charge of various functions in SHP allowed for a participatory meta-evaluation, enriching the results obtained²¹.

An evaluation needs to be useful to those who commission and are interested in it²². Throughout both evaluation processes, evaluators state they have always carried this underlying concern. For example, given the context disparity between the time NHDP I was drawn up and its implementation evaluation, one of the evaluators referred that *“The great challenge was to transform undesirable assessed situations into contributions to the growth of that country”*.

In this meta-evaluation, the checking of evaluation data further use for both processes under analysis met one of the meta-evaluation greatest challenges, which was to verify to what extent the results were used, considering evaluation the more useful the more used it is²².

In RGB, it can be noted that the evaluation of the national health strategy implementation described in NHDP is one of SHP steps, in other words, the first step in the planning process.

Being aware that evaluation cannot be exempted from development agendas^{26,27}, the evaluation process of NHDP implementation as a planning step was, according to interviewed evaluators, *“widely discussed with partners”*. As for people who lived it personally, despite the contextual demands marked by the political instability existing at the time evaluations under analysis were carried out, *“these were very interesting works, because they were done respecting a formative perspective”*. As one assessor points out, referring to NHDP I implementation evaluation:

We, the working team, even discussed how the report should be drafted so that it would not be interpreted as a sanction for what had not been accomplished, but, rather, as a stimulus, a lever, so to, recognizing what had not been done, we could perceive what it would take to accomplish in the future.

Evaluations are an integral part of a process that, although not showing

significant progress in specific indicators, given that the implementation degree of strategies is small, exemplifies an SHP exercise in a FS, allowing the path for a convergence of visions.

Specificity as a differentiating element - the importance of the Guinean context

One of the interviewees synthesizes a key idea of this meta-evaluation as to the specificity of evaluations analyzed as an integral part of SHP process in RGB:

This evaluation and planning process was responsible for the creation of a planning culture in Minsap. We can question the reason for this ever-present desire to plan also at the level of global health in the country as at the vertical programs, and how it was possible to plant this culture in a scenario of contextual instability as RGB'. Eventually, due to the influence of partners, mainly funders, of whom there is so strong dependence. What is certain is that RGB, at its health sector level, may not detain any competence for planning and implementation processes, but these have a strong meaning.

The fragility of the State in RGB, with all that it implies in terms of the country context, emerges as a pattern of specificity^{18,25}, i.e., the fact that RGB presents FS characteristics and that the context analysis has been widely considered in the evaluations carried out allows to characterize them as holders of merit and value, since they are adequate evaluations and respond to the information needs of stakeholders. The broad consideration of the Guinean context in carrying out the evaluations under analysis also leads to think over the adaptation of evaluation methodologies to the context where it is made.

Well planned and executed evaluations are particularly important in countries with fragile institutions and vulnerable populations²⁸.

Since evaluation is a challenging and exciting endeavor, it should create trustworthy and useful knowledge by means of credible and insightful practices²⁸. However, a single recipe for evaluation is impracticable – many evaluations' principles, assumptions and practices developed in highly developed countries are considered inadequate for less developed contexts²⁹. At the global level, a growing concern exist that an evaluation standardized approach to the western evaluation model is not always appropriate in distinct cultural and developing contexts²⁹.

As for developing countries in particular, there is a preponderance of external actors, such as the donor community, in formalizing the evaluation practice. These donor-driven approaches cause dissatisfaction and lead to think over the need to suit and adjust evaluation strategies to different sociocultural, political, economic and ecological contexts²⁷. According to the same authors, such thinking is motivated by the assumption that methodology is context-sensitive.

As far as the African continent is concerned, it is noted that, until the 1980s, evaluation was largely driven by international actors, and that, even today, evaluations in Africa are mainly commissioned by non-African stakeholders, comprised in its majority by international donors or development agencies that administer or fund development programs on the continent²⁹.

Although the desire to confer cultural relevance to evaluation is currently at the center of the discourse on evaluation, global attempts to deploy culturally sensitive practices have yet to integrate African voices. Given the fact that the current western evaluation paradigm is not always applicable to the African context, arguments emerge inside the evaluators' community in favor of an Africa-oriented evaluation paradigm that better suits African conditions, cultures and institutions²⁹.

This thinking challenges those who

believe in evaluation processes guided solely by quantitative indicators that neglect what is less tangible. On the other hand, it reinforces the need for institutional frameworks that ease participatory approaches and recognize the value systems that support evaluation and call upon States to apply evaluation to improve the nature of their governance approaches²⁸. The community of evaluators advocates agreements between government, civil society and business to allow the thinking over the merit and value of evaluations and promote their use²⁸.

In short, it is urgent to consider the context, and, in it, the cultures of peoples as an undeniably integrated part of the various evaluation contexts, thus, an integral part of the evaluation. Culture is present in the evaluation, not only in contexts in which programs are deployed but also in their projects, and in the approach or methods that evaluators choose to adopt³⁰.

Conclusion

A meta-evaluation was carried out due to the existence of evaluation data on the implementation of the national strategic orientation documents for the health sector in RGB, the so called NHDP. This work provided for conclusions about the methodological and procedural quality of each evaluation examined, thus contributing to future evaluation so to increasingly fulfill the excellence of procedures. Major contributions, however, reflect the clear conclusion of the relation between evaluation and planning and with the specificity of these evaluations, conducted in an African State classified as fragile.

This work leaves the challenge of deepening the SHP theme in RGB, reading its history in the context of FSs. On the other hand, the fact that evaluation data are used in planning provokes the curiosity for specific processes that emerge from evaluations and planning, as is the case, for example, of

health human resources training.

NHDP created a Minsap 'planning culture' largely motivated by the sector's funders. These documents have turned out to be the benchmark in instability scenarios, demonstrating that they are upstream of those who govern, resisting adversity, and addressing the sustainability challenges of the RGB health sector.

It should be emphasized that, probably because of evaluators' wide knowledge about RGB reality, evaluations fulfilled the current challenge of introducing in the analysis the issue concerning contexts and cultural spaces in which the evaluation research takes place, taking them as an integral part of the evaluation methodological definition.

As a warning to future processes in which this is not intuitive, we suggest to think over the need to look at evaluations as complex undertakings deeply influenced both by the context and people involved; combining the objectives of an evaluation with actual needs and priorities of stakeholders is essential to ensure the use of evaluation results to support the decision-making.

Authors

Guerreiro CS substantially contributed to the design, planning, analysis and data interpretation, and to the draft and final version of the manuscript. Ferrinho P contributed to the design, planning, content review, and participated in the approval of the final version of the manuscript. Hartz Z substantially contributed to the design, planning, content review, and participated in the approval of the final version of the manuscript.

Acknowledgment

We thank Fernando Cupertino for the valuable collaboration with the orthographic review of the paper. ■

References

1. Guiné Bissau. Ministério da Educação. Relatório do estado do Sistema Educativo para a reconstrução da escola da Guiné-Bissau sobre novas bases. 2015.
2. Messner JJ. Fragile States Index 2015. Washington: The Fund for Peace; 2015.
3. Newbrander W, Waldman R, Shepherd-Banigan M. Rebuilding and strengthening health systems and providing basic health services in fragile states. *Disasters*. 2011; 35:639-660.
4. McPake B, Witter S, Ssali S, et al. Ebola in the context of conflict affected states and health systems: case studies of Northern Uganda and Sierra Leone. *Confl Health*. 2015; 9:23-32.
5. World Health Organization. The impact of global health initiatives on health systems [internet]. Geneva: WHO; c2018 [acesso em 2018 ago 23]. Disponível em: <http://www.who.int/alliance-hpsr/researchsynthesis/project3/en/>.
6. Guiné-Bissau. Ministério da Saúde Pública. Plano Nacional de Desenvolvimento Sanitário II 2008-2017. Guiné-Bissau: Ministério da Saúde Pública; 2008.
7. Organização Mundial da Saúde. Estratégia de cooperação da OMS com os países, 2009-2013: Guiné-Bissau. República do Congo: OMS; 2008.
8. Guiné-Bissau. Ministério da Saúde Pública. Plano Nacional de Desenvolvimento Sanitário 1998-2002. Guiné-Bissau: Ministério da Saúde Pública; 1998.
9. 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.
10. Guiné-Bissau. Ministério da Economia, do Plano e da Integração Regional. Deuxième Document de Stratégie Nationale pour la Réduction de la Pauvreté-DENARP II 2011-2015. Guiné-Bissau: Ministério da Economia, do Plano e da Integração Regional; 2011.
11. Ferrinho P. Subsídios para a Revisão do Plano Nacional de Desenvolvimento Sanitário 2008-2017 até 2020 - apreciação da sua implementação em Agosto de 2015 e contribuição para um roteiro para a sua revisão. Guiné-Bissau; 2015.
12. European External Action Service. Terra Ranka: A Fresh Start for Guinea-Bissau. Guiné-Bissau: EEAS; 2015 [acesso em 2017 out 26]. Disponível em: https://eeas.europa.eu/headquarters/headquarters-homepage_en/2105/Terra.
13. Hartz ZMA. Meta-avaliação da gestão em saúde: desafios para uma nova saúde pública. *Ciênc Saúde Colet*. 2012; 17(4):832-834.
14. Bardin L. Análise de Conteúdo. 5. ed. Lisboa: Edições 70; 2008.
15. Flick U. Métodos qualitativos na investigação científica. Lisboa: Monitor; 2005.
16. Marconi M, Lakatos E. Fundamentos de metodologia científica. 2. ed. São Paulo: Atlas; 2007.
17. Samico I, Felisberto E, Figueiró AC, et al. Avaliação em Saúde: bases conceituais e operacionais. Rio de Janeiro: Medbook; 2010.
18. Hartz ZMA. Principles and standards in metaevaluation: guidelines for health programs. *Ciênc Saúde Colet*. 2006; 11(3):733-738.
19. Hartz Z, Goldberg C, Figueiro AC, et al. Multi-strategy in the Evaluation of Health Promotion Community Interventions: An Indicator of Quality. In: Potvin L, McQueen DV, Hall M, et al. *Health Promotion Evaluation Practices in the Americas*. New York: Springer; 2008. p. 253-267.

20. 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.
 21. Furtado JP, Laperrière H. Parâmetros e paradigmas em meta-avaliação: uma revisão exploratória e reflexiva. *Ciênc Saúde Colet.* 2012; 17(3):695-705.
 22. Elliot LG. Metaevaluation: from approaches to possibilities of application. *Aval Pol Públ Educ.* 2011 out-dez; 19(73):941-964.
 23. Yarbrough DB, Shulha LM, Hopson RK, et al. *The Program Evaluation Standards: a guide for evaluators and evaluatin users* [internet]. 3. ed. Thousand Oaks: Sage; 2011 [acesso em 2017 nov 15]. Disponível em: <http://www.jcsee.org/program-evaluation-standards-statements>.
 24. Hartz Z, Felisberto E. *Meta-avaliação da Atenção Básica à Saúde: teoria e prática*. Rio de Janeiro: Fiocruz; 2008.
 25. Brousselle A, Champagne F, Contandriopoulos AP, et al. *Avaliação: conceitos e métodos*. Rio de Janeiro: Fiocruz; 2011.
 26. Craveiro I, Hartz Z. A equidade na investigação avaliativa com foco na cooperação em Saúde para o desenvolvimento. *An Inst Hig Med Trop.* 2017; 16(supl):S31-S38.
 27. Carden F, Alkin MC. *Evaluation Roots: An International Perspective*. *J Multidiscip Eval.* 2012; 8:102-118.
 28. Fanie Cloete, Babette Rabie, Christo de Coning. *Evaluation Management in South Africa and Africa*. [Sem local]: SUN MeDIA Stellenbosch; 2014.
 29. Cloete F. Developing an Africa-rooted programme evaluation approach. *African J Public Affairs* [internet]. 2016; 9(4):55-70 [acesso em 2017 dez 6]. Disponível em: <http://repository.up.ac.za/handle/2263/59022>.
 30. Chilisa B, Major TE, Gaotlhobogwe M, et al. Decolonizing and Indigenizing Evaluation Practice in Africa: Toward African Relational Evaluation Approaches. *Can J Program Eval.* 2016; 30(3):313-328.
-
- Received on 03/02/2017
Approved on 08/16/2018
Conflict of interest: non-existent
Financial support: non-existent