



I Gede Sudi Adnyana

Badan Pemeriksa Keuangan Republik Indonesia

Jalan Gatot Subroto Kav. 31, Jakarta Pusat

igede.adnyana@bpk.go.id



## **MEMAHAMI METODE PEMERIKSAAN UNTUK MEMBANTU MERENCANAKAN EVALUASI PROGRAM: BERBAGI PENGALAMAN YANG DITERAPKAN GAO ATAS PROGRAM JAMINAN KESEHATAN NASIONAL**

## **UNDERSTANDING AUDIT METHODS TO ASSIST PLANNING ON PROGRAM EVALUATION: SHARING GAO'S EXPERIENCE ON NATIONAL HEALTH SECURITY PROGRAM**

### **ABSTRAK/ABSTRACT**

BPK mempunyai mandat melaksanakan pemeriksaan kinerja untuk menguji apakah unit kerja pemerintah telah melaksanakan program secara ekonomis, efisien dan efektif. Namun demikian, metode pemeriksaan atas program pemerintah yang dirancang pada tahap perencanaan oleh tim pemeriksa belum sepenuhnya menekankan pada salah satu atau kombinasi dari empat jenis evaluasi yang ada, seperti evaluasi proses atau implementasi, evaluasi hasil, evaluasi dampak, serta analisis atas efektivitas beban dan beban-manfaat. Pada dasarnya, evaluasi tersebut dapat diidentifikasi dengan menggunakan ukuran kinerja program, beserta informasi lainnya untuk mengetahui manfaat dari suatu program atau bagaimana meningkatkannya. Topik yang dipilih adalah Program JKN dengan perencanaan evaluasi menggunakan pengalaman GAO, yang sudah dipraktikkan secara empiris dan diobservasi oleh penulis saat mengikuti program fellowship. Penulis membahas tentang evaluasi terkait fragmentasi, overlap, dan duplikasi (FOD) serta panduan umum untuk menjawab pertanyaan pemeriksaan yang berasal dari evaluasi tahapan Program JKN, analisis jenis pertanyaan pemeriksaan, serta tantangan dalam merancang evaluasi. Pada akhirnya, BPK dapat lebih meningkatkan metode dan desain rencana pemeriksaan kinerjanya terkait evaluasi program dengan mempertimbangkan variabel-variabel dimaksud yang telah menjadi penerapan terbaik.

#### **KATA KUNCI:**

kinerja; evaluasi program; jaminan kesehatan.

#### **SEJARAH ARTIKEL:**

Diterima pertama: April 2017

Dinyatakan dapat dimuat : Juni 2017

BPK has a mandate to conduct performance audit for examining whether government agencies have implemented their critical programs in the possible aspects of efficiency, effectiveness or economy. Nevertheless, the current audit methods for those government programs designed at the planning stages by audit teams have not been fully emphasized on one or a combination of four main types of evaluation; outcome evaluation; impact evaluation; and cost-benefit and cost-effectiveness analyses. They can be identified by using measures of program performance, along with other information, to study the benefits of a program or how to improve it. The program chosen is National Health Security (NHS) Program, evaluated by using GAO's experience which were empirically implemented and observed by the author during the fellowship program in the Agency. This study discusses exposure of FOD and a general guide to address researchable questions derived from evaluation of NHS Program stage; analysis of researchable question types; and some challenges at designing evaluations. Further, BPK could improve methods and design of its performance audit plan as regards program evaluation by considering such variables that have been applied in their best practices.

#### **KEYWORDS:**

Performance; program evaluation; health security.

## INTRODUCTION

The Audit Board (BPK) as the Supreme Audit Institution (SAI) of Indonesia has an interest to fulfill stakeholders' expectation<sup>1</sup> that government programs have already been directed to attain the public welfare. In these circumstances, BPK has a mandate of doing performance audit to examine if the government agencies have run their critical programs in the possible aspects of efficiency, effectiveness and economy to give optimal output, outcome, and impact for the benefit of Indonesian people (BPK, 2007).

To carry out such mandate, BPK has implemented quality assurance (QA) of performance audit activities including planning, executing, and reporting phases based on upholding core values of independence, integrity and professionalism (BPK, 2009). In that context, this paper will mainly focus on program evaluation at the planning and design phase that critically becomes a part of quality control (QC) elements<sup>2</sup> in the QA mechanism. The QA scope here is to ensure how BPK auditors implement QC to carry out appropriate planning and design to provide better performance audit reports over critical government programs that have significant influence against stakeholders' expectation.

In relation to the performance audit mandate, BPK has a legal basis to enforce its authority to audit the management and accountability of public finance. The position of BPK is vested in the articles 23E, 23F, 23G of the 1945 Constitution; and Act 15/2006 that regulates the Audit Board of the Republic of Indonesia. In performing its duties and authorities,

BPK is also supported by a set of laws in the field of public finance management, namely Act 17/2003 regarding Public Finance, Act 1/2004 regarding Public Treasury, and Act 15/2004 regarding Audit of Management and Accountability of Public Finance. Those omnibus regulations affirm the position and role of BPK as a free and independent public finance audit institution, one of those regards, to conduct the performance audits<sup>3</sup>.

BPK currently has its 2016-2020 Strategic Plan. One of its strategic goals requires BPK to increase the benefit of audits so as to drive prudent public administration to achieve state goals. The other strategic goal also includes the effort through the paradigm of insight and foresight to improve the quality of audit services so as to drive trusted public administration to achieve state goals. Both simultaneously will trigger good audit management in order to provide quality audit results that are in accordance with stakeholders' needs and their making decisions (BPK, 2015). In this regard, the strategic plan covers the capability of performance audit in twelve audit themes to evaluate whether all focused programs have already been directed to attain the public welfare (BPK, 2016). Thus, the performance audit function in health services<sup>4</sup> especially related to National Health Security (NHS) will be paramount in each BPK Annual Plan.

## METHODS

### Underlying Questions

Based on the background information above, this paper subsequently raises three possible

<sup>1</sup> Expectation of the external parties that have political, economic, and social interests, i.e. the Parliament, NGOs, other government agencies, private sector, and the Indonesian people.

<sup>2</sup> One of nine QC elements is Audit Performance as described in the Implementation Guidelines of BPK Quality Assurance System that contains a variety of procedures to review the design and implementation of quality controls.

<sup>3</sup> The omnibus regulations provide authority to BPK to conduct financial, performance, and special purpose audits.

<sup>4</sup> The 2016-2020 Strategic Plan defines health services as one of the audit themes that comprise i.e.: state economy and finances, education, population & family planning, mental & character, food availability, energy & electrical power availability, marine & maritime, regional development, equitable development, security order, good governance & bureaucratic reform.

underlying questions to be considered:

- How should BPK enhance its capability to conduct performance audits mainly during planning and design in relation to quality assurance framework?
- How might BPK adopt program evaluation methods that GAO systematically implements using research methods to collect and analyse data to assess how well programs work?
- To what extent will BPK add value to audited agencies and to public expectation by conducting a performance audit on one example of NHS Programs, which would result in public welfare improvement?

To illustrate how to analyse and potentially adopt applicable program evaluation methodologies, one of two potential topics under the NHS major programs<sup>5</sup> is used to demonstrate their complexity in terms of structure, magnitude, risk, financial significance, sensitiveness, and impact; that is “*management improvement of health insurance in the form of refinement and coordination of benefit packages, provider incentives, quality control and cost of services*”.

The particular potential topic is analyzed using six types of researchable questions aligned with the program stage evaluation to generate key areas. From GAO perspectives, the potential topic could be the base of congressional requests<sup>6</sup>; however, from BPK mandate, it results from the board’s engagement discretion and designation to BPK’s audit directors to choose one over various audit topics in the health services field outlined in the Annual Plan. The different and similar practices of

conducting performance audit at both SAIs would be described implicitly in the content sections. To elaborate both SAIs’ perspectives, the choice of one key area using GAO planning and design phase will become a potential audit objective i.e. “*the extent to which the agencies improve health insurance management in the form of refinement and coordination of benefit packages to end-users/beneficiaries.*”

Next, the potential audit objective subsequently results in researchable questions for audit, which can be addressed using program evaluation by means of fragmentation, overlap, and duplication (FOD) guide (GAO, 2015a). This guide developed using the notion of program logic proximity<sup>7</sup> helps the auditors identify and evaluate instances of fragmentation, overlap, and duplication, also identify options to reduce or better manage associated negative effects. It has four main steps: identify FOD; identify the potential effects of FOD; validate effects, and assess and compare agencies’ programs; as well as identify options to increase efficiency, and reduce or better manage FOD.

### **Background Information of NHS Programs**

One of Indonesian Government priorities described in the National Medium Term Development Plans (NMTDP) 2015-2019 is to improve public welfare through health care program (Sekretariat Kabinet, 2015). It is mentioned specifically in the National Development Agenda Number 6.5.3 under the title of “Health Development: Implementation of Healthy Indonesia Program”. The agenda is separated into two main sections namely: “goals” and “policy direction and strategies”.

<sup>5</sup> The two examples are extracted from the National Development Agenda 2015-2019 Number 6.5.3 under the title of “Health Development: Implementation of Healthy Indonesia Program”, in which the detail is immediately followed in the background information of NHS Programs and Table 1.

<sup>6</sup> 90% of GAO work is based on congressional requests.

<sup>7</sup> See Appendix 1 Program Logic and Stages of Program as an illustration.

The intended goal achievement in NMTDP 2015-2019 is to elevate people's health and nutritional status through health effort and community empowerment that is supported by financial protection and health services equalization. The main goals are improvements in the aspects of:

1. health and nutritional status of mothers and children
2. disease control
3. access and the quality of basic health services and referrals especially in border, marginal and remote areas
4. coverage of universal health services by means of Health Indonesia Card (HIC) and quality of health security management
5. fulfillment of needs of health workers, medicine, and vaccine; as well as
6. health system responsivity.

Meanwhile, the policy and strategies aim at improving people's health and nutritional development at all life cycle consisting of individual, family or society. The reformation mainly focuses on strengthening quality primary health care regarding all of the previously intended improvements. Indonesian Health Card (IHC) becomes one of main means to encourage health sector reformation in order to achieve optimal health service including promotive and preventive efforts.

There are ten<sup>8</sup> parameters of policy and strategies that belong to the agenda of "Health Development: Implementation of Healthy Indonesia Program", which one of those is in significant relation to the substance of this

<sup>8</sup> The other nine parameters in sequence are to enhance 1) fulfillment of quality health service access for mothers, children, teenagers, and the elderly; 2) public nutrition; 3) disease control and environmental sanitation; 5) quality basic health service access; 6) quality referral health services; 7) availability, spread, and quality of human resources in health; 8) availability, affordability, equalization, and quality of pharmacy, and medical devices; 9) food and drug surveillance; 10) health promotion and community empowerment.

paper namely Parameter Number 4: to stabilize the implementation of NHS through seven efforts i.e.:

1. Increase in the number of membership coverage of Healthy Indonesia Card;
2. Increase in the number of health care facilities that become standard-compliant service providers, among other things, through cooperation between the government and private service providers;
3. Management Improvement of health insurance in the form of refinement and coordination of benefits package, provider incentives, quality control and cost of services, improvement of financial system accountability, development of health technology assessment, and development of integrated monitoring and evaluation system;

The effort number 3 would hypothetically become the illustration of key areas among potential audit topics based on BPK's consideration on selection factors<sup>9</sup> such as risk management, impact; significance and auditability (see also in table 1).

4. Completion of payment systems for strengthening primary health care, maternal and child health, health workers' incentives in remote and border areas including the islands; and improvement of individuals' promotive and preventive efforts
5. Development of various regulations including guidelines and standards of health care;
6. Increase of institutional capacity to support quality of service; and
7. Financial development of health services in public private partnership.

<sup>9</sup> The process of applying selection factors to "potential areas" is not shown here for the succinct purpose of this paper.

**Table 1.** An Illustration of Potential Audit Topics throughout Five-Year Period Based on Effort Number 3 of Seven Efforts in the Background Information of National Health Security (NHS) Programs

Examples of "Potential Audit Topics"	Performance Audit to Evaluate Effectiveness or Impact of Activities/Programs			
	2017	2018	2019	2020
1) Improvement of health insurance management in the form of refinement and coordination of benefits package, service provider incentives, quality control and cost of services	Improvement of health insurance management in the form of refinement and coordination of benefit package	Improvement of management in service provider incentives, quality control and cost of services	the impact of 2017 and/or previous years' programs	the impact of 2018-19 programs
2) Improvement of financial system accountability, development of health technology assessment, and development of integrated monitoring and evaluation system			improvement of financing system accountability, and development of health technology assessment	development of integrated monitoring and evaluation system

Those examples contain critical points that auditors would be able to evaluate the program achievement by means of selecting and determining appropriate performance audit methodology, which emphasizes on planning and design in order to possibly draw potential conclusion that can be in forms of lesson learned and/or implication to respond intended audit objectives.

**GAO’s Study on Performance Measures<sup>10</sup> and Program Evaluation**

GAO describes performance measurement as continual monitoring and reporting of program accomplishments, particularly progress against pre-determined goals. It is typically carried out by management of program or agency. The term “program” could be any function, project, activity, or policy, which respectively has an recognizable purpose or set of objectives.

Meanwhile, an approach to do performance audit of a particular program can also

be interchangeably termed as a program evaluation. It is defined as a discrete systematic study performed periodically or on an ad-hoc basis to assess if a program is working well. The evaluation is often undertaken by experts external to the program. The experts could be either inside or outside the agency, or even program managers. A program evaluation typically examines achievement of program objectives in the context of program performance. It can also be in the context in which it occurs. Four main types of evaluation can be identified such as evaluation of process and/or implementation; outcome evaluation; impact evaluation; and cost-benefit and cost-effectiveness analyses. They all use measures of program performance, along with logical assumptions and other relevant information, to observe the benefits of a program or possible ways to improve it. The objective of program evaluations is to assess if program expectations have been met (GAO,2012). It generally does not assess a program’s or agency’s compliance with legal requirements, to which financial or performance audits lead to some extent.

There is relationship between program evaluation and performance measurement

<sup>10</sup> Indicators used to assess how well an organization is achieving its desired objectives through processes; and related to results (outputs or outcomes) or determinants of the results (inputs such as quality, flexibility, resource utilization, and innovation).

(GAO, 2011b). A program evaluation typically examines a wider range of information on program performance and its context, which is relatively not feasible to monitor it on an ongoing basis. On the other hand, performance measurement expressed toward measurable performance standards focuses on whether a program has achieved its objectives (Performance Measure, n.d.).

In aspects of different intended focus, evaluation can observe aspects of program operations i.e. a process evaluation or factors that might hinder and/or contribute to its success in the program environment. So that, it will help clarify the connections between or among program inputs, activities, outputs, and outcomes. As an alternative, it may also assess the program's effects beyond the program intended objectives. In respect of the program's net impact, evaluation could be used to estimate what would have occurred in the absence of the program. Further, program evaluations may analytically compare the effectiveness of alternative programs with a similar objective.

In aspects of different use, both performance measurement and program evaluation aim to assist policy decisions and resource allocation to enhance program effectiveness and service delivery. Due to its ongoing nature, performance measurement can play a role as an early warning system to management and as a vehicle for improving accountability to the public. On the other hand, a program evaluation examines program performance more profoundly. Its context enables the evaluator to overall assess whether the program works, and to identify adjustments that may improve its results.

Related to performance measures, four types of program evaluation cover: 1) process or implementation evaluation that measures the extent to which a program is working as intended. The assessment focuses on program activities' conformance to statutory and regulatory requirements, program design, professional standards, and stakeholders' expectations; 2)

outcome evaluation that measures the extent to which a program achieves its outcome-oriented objectives. Its assessment focuses on outputs and outcomes to determine program effectiveness, which may include unintended effects. It may also assess program process to recognize how outcomes are produced; 3) impact evaluation that measures the net effect of a program by comparing program outcomes with an estimate of what would have happened in the absence of the program. This form of evaluation is performed when external factors are known to influence the program's outcomes so as to isolate the program's contribution to achievement of its objectives; and 4) cost-benefit and cost effectiveness analyses, which are regarded as one type of program evaluations. They compare a program's outputs or outcomes with its input namely the costs of resources disbursed to produce them. Cost-effectiveness analysis evaluates the cost of meeting a single goal or objective and seeks efforts to identify the least costly alternatives. Cost-benefit analysis commonly articulated in monetary unit has a purpose to determine all relevant costs and benefits.

## RESULT AND DISCUSSION

### **Practices of Performance Audit and Program Evaluation Methods**

At the acceptance phase, GAO uses a risk-based management approach to determine all engagements from initiation of work through product issuance comply with the core values and professional standards (GAO, 2008). Furthermore, risk-based management involves managing the costs of achieving quality and making the level of resources investment to develop quality throughout an engagement commensurate with risks. This approach also establishes clear responsibility and accountability at all levels for quality throughout an engagement, including engagement design, staffing, internal stakeholder involvement,

**Table 2.** Performance Audit Objectives at Acceptance Phase

Possible Audit Year	Potential Audit Objectives based on Requests or BPK-Determined Key Areas	Potential Approaches to Program Evaluation related to Potential Audit Objectives	GAO’s Risk Based Management Approach
2017	<ul style="list-style-type: none"> <li>the extent to which the agencies improve health insurance management in the form of refinement and coordination of benefit packages to end-users/beneficiaries.</li> </ul>	<ul style="list-style-type: none"> <li>Process or Implementation Evaluation</li> <li>Outcome Evaluation</li> </ul>	<ul style="list-style-type: none"> <li>A risk factor—low, medium, or high—is assigned to the engagement that will, among other things, determine the level of product review and executive-level involvement throughout the engagement. Factors such as the human capital (mission team), costs, complexity, or potential controversy of the work and related potential access to records concerns are considered in the risk determination.</li> <li>Statutory basis for access to Medicare records/sensitive information that can be impediment to engagement</li> <li>Limitations on the use of data and confidentiality agreements</li> <li>Potential limitation of audit authority</li> </ul>
2018	<ul style="list-style-type: none"> <li>the extent to which the agencies manage service provider incentives, quality control and cost of services for customer benefit</li> </ul>	<ul style="list-style-type: none"> <li>Cost-Benefit and Cost Effectiveness Analyses</li> <li>Outcome Evaluation</li> </ul>	
2019	<ul style="list-style-type: none"> <li>the extent to which the agencies improve financing system accountability, and development of health technology assessment<sup>11</sup> for customer benefit.</li> <li>assessing 2017 and/or previous years’ program impact to NHS beneficiaries after two years’ implementation</li> </ul>	<ul style="list-style-type: none"> <li>Process or Implementation Evaluation</li> </ul>	
2020	<ul style="list-style-type: none"> <li>the extent to which the agencies develop integrated monitoring and evaluation system for agencies’ and public benefits</li> <li>assessing 2017, 2018, and/or 2019 program impact to beneficiaries of after three years’ implementation</li> </ul>	<ul style="list-style-type: none"> <li>Process or Implementation Evaluation</li> </ul>	

Source: discussions with relevant parties as well as author’s assumptions and analysis

message development, and product review.

Recalling the illustration to select one potential topic introduced at the beginning of this paper and related directly to table 1, the following table 2 precedes the decision process to select the potential audit objective, which much more elaborates GAO perspective at the Acceptance Phase.

**Planning and design phase**

Next, the planning and design phase follows immediately after the acceptance phase. After

work is initiated, GAO’s senior managers periodically reassess the need for continued high-level involvement in the work at the bi-weekly engagement review meeting, which is structured according to GAO’s strategic goals and continuation of risk based management approach. The focus is to provide medium and high risk engagements the appropriate level of scrutiny to develop products that will meet Government Auditing Standards and GAO’s quality standards and will be cost-effective, timely, fair, balanced, and objective (GAO, 2011a).

<sup>11</sup> WHO of UN defines it a multidisciplinary process to evaluate the social, economic, organizational and ethical issues of a health intervention or health technology as the application of organized knowledge and skills in the form of medicines, medical devices, vaccines, procedures and systems developed to solve a health problem and improve quality of life.

For medium and high risk engagements, the initial focus is on the engagement’s design, scope, and methodology. As the engagement moves through the data gathering and analysis

phase, the focus is on the extent to which the objectives are being achieved and that the product will meet the needs of the client.

The planning and design phase consists of assigning staff, engagement planning, consultation with experts and specialist, as well as stakeholder involvement. One of important elements in the engagement planning is developing audit plan that includes the design matrix describing the audit objectives, scope, and methodology. It is subject to update as needed. The level of detail in the design matrix or paper may vary depending on the nature of the audit, but more detail may be needed for complex or sensitive topics, or topics that are new to GAO; staff that are new to an issue; risky, innovative methodologies; and audits using data with unknown reliability.

### **Six types of researchable questions**

In constructing a design matrix, researchable questions become key drivers to determine the content of the remaining elements or columns across the matrix, so that analysts<sup>12</sup> are advised to also consider six types of researchable questions (GAO, 2015b) such as: 1) descriptive questions provide reliable information about a condition without assessing the condition against criteria; 2) compliance questions correspond to conformity with criteria established by contracts, provisions, agreements, bylaws, regulations, and other requirements. Those sets of law could affect acquisition, protection, use, and disposition of the entity's resources. They also influence timeliness, cost, quantity, and quality, all of which are related to production and delivery of entity's outputs; 3) internal control questions relate to determine the cause of unsatisfactory program performance. It is due to internal

control involves the processes and procedures for planning, organizing, actuating, and controlling program operations. It also includes management's system for measuring, reporting and monitoring program performance; 4) economy and efficiency questions focus on the terms of economy and efficiency. They seem combined in formulating pertaining questions since both are closely intertwined to address the costs and resources used to achieve program results; 5) program effectiveness questions focus on program effectiveness and results. They attempt to measure the manner and extent to which a program is achieving its goals and objectives. They may also observe the quality of program implementation; 6) prospective analysis questions give analysis or implication about information that is based on logical assumptions about events that may occur in the future. They may include analysis of possible actions of audited entities to possibly respond to future events.

### **Critical elements of the design matrix applicable to NHS Programs**

Recalling table 2 to illustrate the application of the design matrix concisely, one potential audit objective i.e. "the extent to which the agencies improve health insurance management in the form of refinement and coordination of benefit packages to end-users/beneficiaries" is selected as an outline example, in which the remaining ones can adopt the similar pattern into the logic of constructing its respective design matrix. In addition, the selection is potentially relevant to the adjacent year of 2017, and it is also based on the mature stages of NHS as a large program that has already been running for a long time under the different names<sup>13</sup>. Nevertheless, there is a nuance of intersection with the early stage of the program because it

<sup>12</sup> GAO refers analysts as a general term for auditors and the internal stakeholders i.e.: specialists, attorneys, methodologists and others with expertise relevant to the accomplishment of an engagement's objectives.

<sup>13</sup> It relates to Common Evaluation Questions Asked at Different Stages of Program Development in Designing Evaluation (revision), January 2012 (see Appendix 1 for more detail).

is currently under the year-end 2014 elected central government and the NMTDP of 2015-2019. In this context, there are also possibly less large crosscutting programs under the activity of insurance management within the Ministry of Health with other related agencies so that the approaches of process evaluation, outcome evaluation, and impact evaluation within the element of scope and methodology will be discussed subsequently. The next paragraphs will to some extent align the mechanism

of six types of researchable questions with program stage evaluation approach; and focus on the scope and methodology as well as the limitations, which BPK can adopt from GAO.

Since the analysis of one potential audit objective results in various potential researchable questions, the questions related to the process evaluation is relatively more feasible to be measured in that current year compared to those related to the outcome and impact evaluations. It is because longer time is needed

**Table 3.** Potential Researchable Questions Analyzed through Six Types and Compared to Program Stage Evaluation<sup>14</sup>

Potential Researchable Questions using Mechanism of "Six Types"	Potential Considered Researchable Questions based on "Program Stage Evaluation"	Type of Evaluation
<b>Descriptive-type questions:</b>		
<ul style="list-style-type: none"> <li>How many type of insurance are in place related to insurance management activity?</li> <li>What about the differentiation among population to targeted insurance premium?</li> <li>How many low-income people are not covered yet by the insurance administration?</li> </ul>	<ul style="list-style-type: none"> <li>How many low-income people are not covered yet by the insurance administration?</li> </ul>	<ul style="list-style-type: none"> <li>Process Evaluation on early stage of program but close to mature stage in the context of "What progress has been made in implementing changes or new provisions?"</li> </ul>
<b>Compliance-type questions:</b>		
<ul style="list-style-type: none"> <li>How many agencies are involved in the provision and delivery of insurance management according to the NHS Act?</li> <li>To what extent the agencies comply with the NHS Act and other statutes?</li> <li>To what extent do programs like Medicare or Medicaid monitor charges to ensure health care providers adhere to the fee schedule and insurance coverage?</li> </ul>	<ul style="list-style-type: none"> <li>To what extent do programs like Medicare or Medicaid monitor charges to ensure health care providers adhere to the fee schedule and insurance coverage?</li> </ul>	<ul style="list-style-type: none"> <li>Process Evaluation on mature stage of program in the context of "Why is a program no longer obtaining the desired level of outcomes?"</li> </ul>
<b>Internal control questions:</b>		
<ul style="list-style-type: none"> <li>To what extent do the agencies have system in place to monitor the insurance management?</li> <li>To what extent does each agency use a risk based approach to manage and refine the benefit packages?</li> <li>To what extent does each agency ensure that control is in place to allocate limited resources to run the insurance management process?</li> </ul>	<ul style="list-style-type: none"> <li>To what extent does each agency ensure that control is in place to allocate limited resources to run the insurance management process?</li> </ul>	<ul style="list-style-type: none"> <li>Process Evaluation on mature stage of program in the context of "Are program resources being used efficiently?"</li> </ul>

<sup>14</sup> GAO, also emphasizes its audit or engagement function in the aspects of insight, oversight, and foresight.

Potential Researchable Questions using Mechanism of "Six Types"	Potential Considered Researchable Questions based on "Program Stage Evaluation"	Type of Evaluation
<p><b>Economy and efficiency questions:</b></p> <ul style="list-style-type: none"> <li>To what extent coordination and collaboration among the agencies is in place to manage the insurance provision process?</li> <li>To what extent is there overlap or duplication in coordination for insurance management and benefit package provisions among the agencies?</li> <li>What factors explain the variation in cost of patient care among major different type of class of hospitals?</li> </ul>	<ul style="list-style-type: none"> <li>To what extent is there overlap or duplication in coordination for insurance management and benefit package provisions among the agencies?</li> </ul>	<ul style="list-style-type: none"> <li>Process Evaluation on mature stage of program in the context of "Are program resources being used efficiently?"</li> </ul>
<p><b>Program effectiveness questions:</b></p> <ul style="list-style-type: none"> <li>How effectively does the insurance management benefit the end-users?</li> <li>How do the agencies address disadvantage impact of the insurance management?</li> <li>To what extent the end users perceive the benefit packages satisfactorily?</li> </ul>	<ul style="list-style-type: none"> <li>How effectively does the insurance management benefit the end-users?</li> </ul>	<ul style="list-style-type: none"> <li>Outcome Evaluation on mature stage of program in the context of "Are desired program outcomes obtained?"</li> </ul>
<p><b>Prospective analysis questions:</b></p> <ul style="list-style-type: none"> <li>What are strengths and limitations of process of the insurance management?</li> <li>To what extent do the agencies perceive that the existing policy and regulations need to be refined?</li> <li>What are the strengths and limitations for the agencies to coordinate and deliver best services to benefit the end users?</li> </ul>	<ul style="list-style-type: none"> <li>What are the strengths and limitations for the agencies to coordinate and deliver best services to benefit the end users?</li> </ul>	<ul style="list-style-type: none"> <li>(Net) Impact Evaluation on mature stage of program in the context of "Is one approach more effective than another in obtaining the desired outcomes?"</li> </ul>

Source: discussions with relevant parties as well as author's assumptions and analysis

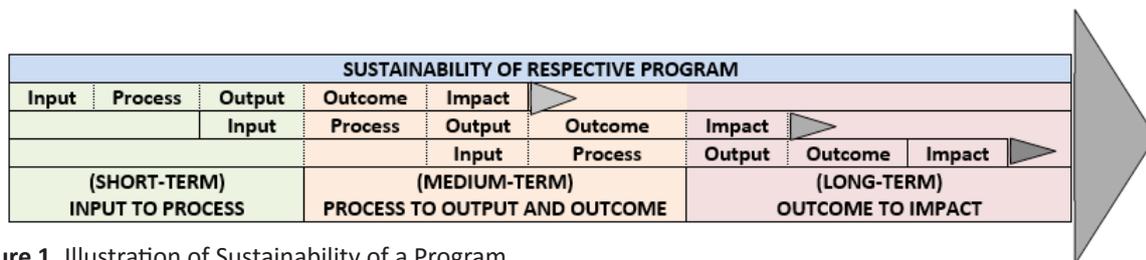


Figure 1. Illustration of Sustainability of a Program

for the outcomes or impacts to show the result; for instance the latter can relatively become the parts of potential researchable questions in the audit year of 2018 or through 2020. However, the analysis really depends on the analysts, evaluators, or managers' perspectives whether the program sustainability is in short, medium, or long-term timeframe, which is always continuing from one term to another. In this case, short term programs may also have their short outcomes or impacts, which can support and influence medium- or long-term ones as

shown in figure 1.

### Consideration of Fragmentation, Overlap, and Duplication (FOD) for Audit Design

This analysis is conducted based on GAO work on FOD to address the audit objective of "the extent to which the agencies improve health insurance management in the form of refinement and coordination of benefit packages to end-users/beneficiaries" (GAO,

2015a). It has three primary researchable questions related to the process evaluation that subsequently leads to outcome evaluation, and impact evaluations i.e.: 1) to what extent is there FOD in coordination for insurance management and benefit package provisions among the agencies? 2) how effectively does the insurance management benefit the end-users while considering the potential existence of FOD? 3) what are the strengths and limitations for the agencies to coordinate and deliver best services to benefit the end users? Those series of questions, especially question 1 and 2, will be discussed in the next paragraph depicting four main steps of conducting FOD review.

### Step 1: identification of FOD in the NHS Program

This step illustrates the possibility of an analysis to identify the FOD among a set of designated programs. In this case, fiscal pressures and limited resources are classical so they both necessitate a continuous review of government programs to ensure their efficiency and effectiveness. Determining whether FOD occurs among programs is an important step to identify opportunities to improve efficiency and effectiveness, which include six procedures:

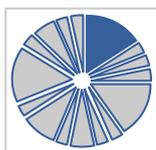
<b>1. Identify an approach for selecting programs for a FOD review:</b>	
Analysts have to opt for one or more approaches or principles to guide in identifying programs to be included in a FOD review. The chosen approaches will affect both the scope of programs covered in the review and the identification and extent of FOD among programs. Potential approaches can comprise as follows:	
<b>Goals and outcomes:</b> NHS Program designed to provide mutual insurance and benefit packages to improve the health and well-being of low income or marginalized patients and to provide affordable cost of health for non-low income patients.	<b>Beneficiaries, users, or other target population:</b> Programs that target individuals and families all experiencing possible adversity to afford high cost of treatment
<b>Key benefits, services, or products:</b> Programs that facilitate easiness for permanent, transitional, and emergency illness; or that provide medical benefits	<b>Administering agencies or agency organizational structure in Indonesian NHS Program context:</b> Programs administered mainly by the Ministry of Health (MH), General Hospitals (GH), Ministry of Coordinator for Human and Cultural Development (MCHCD), Agency of Social Security Administrator in Health (ASSAH), and Agency of Population and Family Planning (APFP); as well as their cooperation with those in local governments such as Health Services (HS) including its Public Health Centers (PHC), Local Hospitals (LH), and Agencies of Woman Empowerment (AWE)
<b>Budget:</b> Programs to address health insurance provisions are identified in budget materials of cross agencies within Indonesian Central and Local Governments	
<b>2. Observe programs to asses a FOD review from previous or current sources of information</b>	
<b>Legal sources,</b> such as legislation chronologies, working groups and committee reports, and regulations on NHS Program	<b>Agency sources</b> in the scope of NHS Program, namely budget documents i.e. requests, proposals, submissions, or justifications; financial statements and reports; performance documents i.e. strategic plans, performance plans, and performance reports; program guidance; internal evaluations; program evaluations; and organizational structure.
<b>Program inventories, catalogues, or databases</b>	
<b>Media sources,</b> including newspapers, social media, television news and the internet	
<b>Non-agency sources,</b> including previous and current BPK reports; inspector general reports; and third-party reports in the scope of law review articles, consultants, working groups, commissions, research groups, and academics.	If required information is nonexistent, unreliable, or limited; analysts will need to carry out <b>original research</b> . Original research may include surveys or interviews to program administrators, beneficiaries, customers, or subject-matter experts; reviews on agency or provider files; or mapping of benefits, services, or processes of NHS Program

**3. Collect background information on the identified programs (closely linked to Procedure 1)**

History and origin of NHS Program	Observations on acts, statutes, and jurisdictions
Functional classification of budget or activity line	Clarifications on appropriation in cross-years or funding allocation
Administering agency on NHS Program; and relevant offices, bureaus, or units within the agency	Analysis on purpose, goals, and activities performed in each agency's official function/structure
Requirements and eligibility of participation	Targeted qualifying beneficiaries or customers

**4. Determine whether FOD exists among the selected NHS Programs:**  
 Analysts should verify if any FOD occurs among the programs after considering current sources of information and performing original research to gather background information on the identified programs. Analysts have to define or use existing definitions of FOD that best meet the needs of their reviews.

Fragmentation indicates to those circumstances in which more than one agency in the central or local government; or more than one function within an agency engage in the same broad area of national need, and opportunities exist to improve service delivery	Overlap arises when multiple agencies or programs have similar goals, engage in similar activities or strategies to achieve them; or they target similar beneficiaries	Duplication signifies when two or more agencies or programs are involved in the same activities or provide the same services to the same beneficiaries
---	--	--



Analysts can elaborate as practical methods as possible to compare among programs i.e. **qualitative approaches**—namely compiling and assessing specific descriptions of similarities and differences among programs—and **quantitative approaches**—namely marking the degree of similarity among programs or classifying observations

**5. Analysts need also measure whether fragmented, overlapping, or duplicative programs are related.**  
 To reveal these connections, analysts can seek advice of current sources of information. Analysts need to perform original research to assess if a program relies on, or is relied upon by, another program or entity. Dependence to each other could be important to achieve intended outcomes. It is also critical to recognize how changes in one program might affect other beneficiaries, programs, and entities.

For example, in **Indonesian NHS context**: Ministry of Health (MH) becomes a leading agency for the health insurance system in which some critical programs are in place such as programs for:

- identifying population of very poor, poor, or low income people or families who are vulnerable to illness and unaffordable to cover the treatment costs
- identifying non-low income people who can support their premium to cover free or small amount of premiums of low income people, a program to gathering such data and information of overall population whether utilizing current data produced by Bureau of Statistics, previous data or conducting new approaches of gathering data or a combination of both, in coordination with Ministry of Coordinator for Human and Cultural Development (MCHCD) who may run the same program;
- analysing amount of premium and cost of treatment in coordination with General Hospitals (GH) and Agency of Social Security Administrator in Health (ASSAH)
- disbursing social assistance in cash to cover treatment cost of low-income women and their children in coordination with Agency of Population and Family Planning (APFP) disbursing grant in cash to cover treatment cost for local people in coordination with Health Services (HS) in local governments and facilitation to use Public Health Centers (PHC) and Local Hospitals (LH).

**6. Validate observations and findings with relevant agencies and related key stakeholders**

Analysts have to confirm the accuracy and completeness of the subject matter information in regards of NHS Program with relevant agencies and related key stakeholders, which include academics, subject-matter experts, and associations. The validation can be undertaken provided compiling a list of programs and identifying any FOD among those programs has finished. For instance, analysts could verify a list of programs to better ascertain the completeness of observations or findings. Their relevant and valid reasoning and propositions result from identification primarily through analyses of agency documents, interviews with agency officials, and surveys to key personnel in the program management.

**Step 2: identification of the potential effects of FOD in the NHS Programs**

This step encapsulates mechanism to identify the potential positive and negative effects of any FOD found in Step 1. The effects can result

from the situation identified, which will help analysts and evaluators determine if there are any potential efforts to improve efficiency, reduce, or better manage the FOD. The step can be assigned into three ensuing procedures:

**1. Identify potential positive and negative effects of the FOD**

Analysts have to develop a comprehensive list of the potential positive and negative effects by exploring further questions such as “is there any evidence of the following assessment area that in respect of.....“

Program Implementation of NHS	Outcomes and Impact of NHS Program	Potential Cost-Effectiveness of NHS Program
<ul style="list-style-type: none"> <li>• program management and agencies work mutually to offer logical and coordinated benefits, products or services?</li> <li>• program management and agencies have strategic and service level agreements in place to enable the achievement of outcomes?</li> <li>• clear roles and responsibilities of related programs and agencies are in place?</li> <li>• the collective programs cover people who might be eligible for benefits, products, or services?</li> <li>• the FOD intentionally planned for the reason to fill a gap, complement an existing program, or try a new method is still necessary or justified?</li> <li>• any unintentional and unplanned FOD has positive effects?</li> </ul>	<ul style="list-style-type: none"> <li>• collective provision of products, services, and benefits aids agencies to fulfil the single and/or shared goals and objectives of their programs?</li> <li>• agencies manage to measure the effort as a whole, if several agencies and programs are working mutually to fulfil shared goals and objectives?</li> <li>• related programs and outcomes are complementary or not operating at cross-purposes or conflicting?</li> <li>• those who are eligible eventually receive products, services, and benefits offered?</li> <li>• a coordinated manner, in terms of recipients not obtaining similar or duplicative products, services, and benefits, etc. from multiple programs, is reinforced?</li> </ul>	<ul style="list-style-type: none"> <li>• economical and efficient aspects to the collective provision of benefits, services, or products are considered carefully?</li> <li>• no reduction in benefits resulting from the current structure of providing benefits, services, or products exists without any reasonable reason for the curtailment?</li> </ul>

**2. Assess if there is a need for further evaluation**

Analysts need to review their lists of potential positive and negative effects of FOD in the context of NHS Program to decide if further evaluation is required.

<p>Analysts have to proceed to Step 3 to further evaluate and compare the performance of the identified programs if they observe potential negative effects or are not sure if the potential effects are positive or negative. Understanding program effectiveness in terms of performance and cost, as well as how programs are interconnected or collaborated can facilitate analysts identify corrective actions to reduce or better manage FOD, which is explained further in <b>Step 4</b>.</p>	<p>Alternatively, analysts may professionally decide that no additional analysis or corrective action is needed to reduce or better manage FOD if they observe all or mostly potential positive effects. In this case, analysts may choose to proceed to <b>Step 4</b> for options to enhance efficiency</p>
--	--

**3. Confirm observations and findings with relevant agencies and related key stakeholders**

<p>Analysts need to confirm and validate their observations and findings with relevant agencies and related key stakeholders after recognizing the potential positive and negative effects of FOD obtained in <b>Step 1</b>. For instance, analysts may wish to interview key agency personnel to confirm positive or negative effects of FOD identified through surveys over beneficiaries</p>	<p><b>Step 3</b> provides perspectives and information on how to use the results of existing or new evaluations to validate the effects identified in <b>Step 2</b>. It also assess and compare programs to determine if they achieve optimal performance and cost-effectiveness</p>
---	--

**Step 3: validation of effects as well as assessment and comparison among agencies' NHS Programs**

This step emphasizes on how to use the results

of existing or new evaluations to validate the effects of FOD identified in Step 2. This step can also be used to assess and compare the programs identified in Step 1 to determine their relative performance and cost-effectiveness.

<b>1. Identify existing evaluations and assess their relevance, validity, and reliability of NHS Program</b>	
<p>If analysts need to assess the purpose, scope, and design of existing evaluations if they observe existing evaluations on the chosen programs. This purpose is to determine if the chosen programs are relevant to the areas in which potential positive or negative effects were identified in Step 2 that involves implementation, outcomes and impact, or cost-effectiveness).</p> <ul style="list-style-type: none"> <li>• Purpose creates the questions that an evaluation will, and will not respond</li> <li>• Scope defines subject matter of the evaluation within particular phases namely as part of a program, an entire program or several programs, as well as the time periods and locations that will be included in the evaluation</li> <li>• Design outlines the information needed to conduct the evaluation, data collection and analysis methods, and limitations to the evaluation.</li> </ul>	<p>Analysts have to assess the reliability and validity of the evaluations considered appropriate to their FOD review. Purposely, all selected evaluations should be assessed against fundamental standards related to research design, conduct, analysis, and reporting. Major weaknesses in these areas or limitations in the availability of program information that affect the reliability or validity of each study's findings must be identified and considered in using the study and placing confidence in the study's findings</p>
<b>2. Perform a new evaluation in case existing evaluations are not available, relevant, or reliable to complement the reviews of analysts regarding NHS Program</b>	
<ul style="list-style-type: none"> <li>• When resources are scarce, the leaders of agencies may also wish to perform evaluations of their programs. The result can be used to justify the continued funding for their programs. To permit drawing valid conclusions about the programs in the future, designing an evaluation needs to involve selecting appropriate measures and comparisons.</li> </ul>	<p>In addition to program performance and cost-effectiveness, analysts may wish to consider further to assess the positive and negative effects of FOD identified in Step 2 when designing a new evaluation.</p>
<b>3. Assess and compare the performance of NHS Program and use evaluations to validate the actual effects of FOD</b>	
<p>Assessment and comparison of the performance and cost-effectiveness of programs can assist analysts determine which programs or aspects of programs are (not) performing well. It can also determine which ones to be recommended for further actions, consolidation, or elimination. The results from new appropriately designed evaluations and the existing ones can help analysts to:</p> <ul style="list-style-type: none"> <li>• assess a single program: how effective and efficient is the program performance? how is the program's desired effect on the target population? what is the cost-per good/service/product delivered by the program?</li> <li>• compare two or more programs: how contrast is the performance and cost-effectiveness of programs?</li> <li>• assess a system (or group) of programs: how efficiently and effectively do these programs support and provide products, services, or benefits collectively? how would costs and benefits change if products, services, or benefits were delivered differently? are there any foregone benefits that could be achieved through integration or better coordination? how contrast is the cost-per- good/service/product delivered by the system of programs against the cost-per-unit of success of the programs individually?</li> <li>• evaluate relationships and interactions between programs: are relationships and interactions necessary, strong, and effective? will they lead to achieve better outcomes?</li> <li>• Likewise, validation over the existence and the extent of positive and negative FOD effects can better arrange recommendations on how to increase efficiency, as well as reduce or better manage FOD.</li> </ul>	
<b>4. Confirm observations and findings with relevant agencies and related key stakeholders</b>	
<p>Analysts need to confirm their observations and findings on the actual effects of FOD, program performance, and cost-effectiveness, with relevant agencies and related key stakeholders. For instance, analysts could interview program administrators and executive function leaders to obtain feedback on their analysis of existing evaluations. Additionally, any limitations or unavailability of program information observed during evaluations could better underline particular areas that need special attention from the agency. Thus, it help develop adequate evidence that analysts could use to better provide recommendations.</p>	

**Step 4: identification of options to enhance efficiency and reduce or better manage FOD in NHS Program**

This step suggests approach to analysts on identifying options to improve efficiency when the effects of FOD are positive, or

reduce or better manage FOD when the effects are negative. Those options provided to policymakers can help relieve some of the fiscal pressures. They may also improve program effectiveness.

<b>1. Identify choices to increase efficiency in NHS Program</b> Efficiency can generally be defined as retaining services or outcomes by consuming fewer resources or improving the quantity and quality of services or outcomes while retaining resources spent. Enhancing efficiency can assist central and local governments better utilize limited resources. Therefore, even if analysts observed mostly positive effects of FOD, they should still try to identify opportunities to enhance efficiency among the programs. To possibly improve efficiency, there are management approaches as follows:		
<b>Restructure of obsolete government organizations and operations</b>	<b>Implementation of process improvement methods and technology improvements</b>	<b>Implementation of a strategic approach to spending</b>
This approach signifies the process of eliminating outdated or ineffective programs and policies. Those programs and policies that remain relevant could be updated and modernized. The process involve improving their targeting and efficiency through several actions. Such actions include redesigning allocation and cost-sharing provisions, consolidating facilities and programs, and streamlining and reengineering operations and processes	This approach points out the process of implementing methods to increase product quality and decrease costs to improve efficiency. Process improvement methods can involve examining processes and systems to identify and correct costly errors, bottlenecks, or duplicative processes. However, maintaining or improving the quality of outputs is still essential. In relation to process improvement, modernizing processes through investments in technology can generate efficiency gains	This approach shows the process of spend analysis to develop a better understanding on agency's spending on goods and services or taking an organization-wide approach to procurement.

<b>2. Identify choices to reduce or better manage FOD that possibly occurs in NHS Program</b> To identify opportunities for efficiency improvement, analysts have to identify options to reduce or eliminate the impact of FOD effects on beneficiaries, customers, and related parties. For that effort, they should also identify the benefits and trade-offs of their proposed recommendations by:	
<ul style="list-style-type: none"> <li>• <b>considering rooms to improve coordination and collaboration within and across agencies</b></li> <li>• further engaging in activities of performance management</li> <li>• taking into account for programs elimination</li> </ul>	<ul style="list-style-type: none"> <li>• possible actions in relation to changes to statute, regulation, or guidance to revise or explicitly define the roles and responsibilities of agencies and program administrators</li> <li>• possibility for consolidating or streamlining programs</li> </ul>

In conjunction to Procedure 2 above, approaches of identification can include some changes by NHS program administrators or agency’s management; new legislation; or an amalgamation of actions. Moreover, potential trade-offs and unintended or even unprecedented consequences may arise for each identified option taken. Consequently, in determining which negative effects of FOD will be eliminated, analysts should also determine which positive effects might be lost. The potential positive effects of options to lessen or better manage FOD are depicted as follows in accordance with each assessment parameter:

Implementation	Outcomes and Impact	Cost-effectiveness
<b>Potential positive effects of suggested options</b>		
<ul style="list-style-type: none"> <li>The programs mutually will cover all who are truly eligible for products, services, or benefits</li> <li>Program management within and across agencies will work collectively to provide logical and coordinated services, products, or benefits</li> <li>Roles and responsibilities of related program and agency will be clear</li> <li>Related programs and agencies will have strategic agreements in place to help achieve outcomes</li> </ul>	<ul style="list-style-type: none"> <li>Those who are truly eligible will receive products, services, or benefits</li> <li>Beneficiaries will receive benefits in a coordinated manner</li> <li>Agencies will be able to measure the entire effort when agencies and program management work collectively to attain assigned goals and objectives</li> </ul>	<ul style="list-style-type: none"> <li>The provision of benefits, services, or products will be economical and efficient</li> <li>Unless reasonable explanation, there is no reduction in benefits</li> </ul>
<b>Questions to be considered with respect to potential trade-offs and unintended or unprecedented consequences</b>		
<ul style="list-style-type: none"> <li>Will there be important changes to program products, services, or benefits?</li> <li>What kinds of legal restrictions may obstruct implementation?</li> <li>What is the agency’s commitment to and capacity for change?</li> <li>To what extent will other programs be influenced?</li> <li>Will beneficiaries know how and where to acquire products, services, or benefits?</li> </ul>	<ul style="list-style-type: none"> <li>To what extent is the possible effect of change on the consistency of implementation, accountability, and performance measurement?</li> <li>Are there any possibilities that benefits will be lost or diminished?</li> <li>To what extent will a loss of resources and expertise affect the ability to attain goals effectively?</li> <li>What is the possible consequence of change on research and innovation?</li> <li>Is there possible that necessary or protective redundancy would be eliminated?</li> </ul>	<ul style="list-style-type: none"> <li>What are the possible savings or efficiencies resulting from the change; and what is the investment needed to implement change?</li> </ul>

**3. Communicate alternatives to improve efficiency and diminish or better govern FOD to policymakers**  
 Analysts should define who will be liable for executing changes to improve efficiency and diminish or better govern FOD to policymakers. In conveying their observations and findings to policymakers, analysts also need to specify the particular actions. That is necessary to be taken to carry out the options, and any related trade-offs.

**The Importance and General Challenges of Iterative Process in Designing Evaluation**

The nature and factors to determine methodology, objectives, and scope of evaluation often overlaps. Since the methodology, objectives, and scope need to be established collectively, they lead to an iterative process in designing an evaluation

plan. Data limitations or new information surrounding the program may occur as the evaluation arrangement is carried out. So that, it has implications for the sufficiency of the initial plans or the feasibility of answering the earliest questions. The ensuing paragraphs will illustrate some concern analysts have to consider when designing their audit designs and eventually conduct them in the field work.

## **Relevance and quality of available data sources**

In choosing sources of evidence to respond to evaluation questions, analysts should assess whether the sources will provide sufficient and appropriate evidence. The process is needed to support findings and conclusions on the evaluation questions. Sufficiency represents the quantity of evidence that is enough to convince a knowledgeable person about the rationality of findings. Equally, appropriateness signifies the relevance, validity, and reliability of the evidence to support the evaluation objectives. The intensity of effort is required to ensure computer-processed data namely agency records are adequately reliable for use. The effort depends on the extent to which the data will be used to support findings and conclusions. It is also influenced by the level of risk or sensitivity related to the analysis and research.

## **Original data collection**

In order to ensure data are sufficient and appropriate, qualified analysts may examine records, actions, or physical conditions, to ensure whether all those satisfy requirements or particular kinds of criteria. When collecting testimonial evidence by means of interviews or surveys, the analyst has to consider whether the interviewee has appropriate knowledge, candid and accurate in expressing their reports or opinions. Additionally, discreet in data collecting instruments, as well as developing and pretesting questionnaire surveys will assist to ascertain the obtained data are necessarily accurate for the purposes of the analyses.

If evaluators aim to aggregate and generalize the results of a sample survey, it is important to ensure that they collect uniform data from every unit in the sample. In this case, sample survey information is usually acquired through self-administered questionnaires or structured interviews. Most of the information will be gathered in close-ended form. It means

that the respondent chooses from responses determined by the interviewer or offered in the questionnaire. Consequently, designing a well consistent set of responses into the data collection process will help establish the uniformity of data across units in the sample.

## **Sample of observations**

Census or a certainty sample refers to collecting and analysing data on the entire population. It makes sense to include all members of a population in a study because the population is small enough and available resources are feasible. Nevertheless, many government programs cannot be researched by means of a census and the evaluator should determine whether to collect data on a probability or nonprobability sample.

In a probability sample also referred to as a statistical or random sample, each member in the population has a known, nonzero chance of being selected. The results of a probability sample can usually be generalized to the population from which the sample was taken. It is encouraged that analysts seek advice from a sampling specialist to identify how large a sample is needed to obtain precise estimates or detect expected effects of a given size. In contrast, a purposive, nonprobability sample refers to each member in the population has an unknown chance of being selected. In this regard, a sample is chosen from knowledge of the population's characteristics or from a subset of a population. Decision to use probability or nonprobability sampling is a key element of the evaluation or audit design that streams from the scope of the researchable questions.

## **A pilot study of data collection**

Conducting a pilot study or a pretest is important before launching complete data collection. The pretest provides evaluators opportunities to refine the design and test the appropriateness, reliability, and availability of proposed data. The evaluators who propose new data collection

could find that a limited exploration of the proposed design may give a useful reality check if one's assumptions hold true. Pretesting the data collection at one or more fields allows the evaluators to confirm that data are available, as well as forms, interview procedures, and means for gathering them are feasible. The process will also provide a prospect to consider if the analysis methodology is appropriate.

### **Communication with experts and specialists**

Specialists paramount in GAO's practices assist the team to develop products that are legally, methodologically, and analytically sound. They also provide engagement teams with the collective competence to perform high-quality work. This will discuss on obtaining and using assistance from experts and specialists in a performance audit, including, among others, attorneys, accountants, actuaries, methodologists, statisticians, data analysts, survey specialists, economists, engineers, program evaluators, and investigators.

Likewise, obtaining Legal Assistance is crucial because government programs are usually created by law and are subject to specific laws and regulations. An understanding of the laws underlying a program can be essential to understand the program itself. Moreover, because SAIs also coordinate with Parliaments, SAIs' products often lead to legislative action or otherwise influence legal developments. Therefore, SAIs' findings, conclusions, and recommendations should be legally sound and reflect the exercise of professional legal judgment in producing quality audit reports.

There is also a potential need to consider forensic audits and investigative services so as to support a performance audit in designing methods of program evaluation. If required, this particular type of services can be useful in relation to investigate the potency of fraud, waste, and abuse, which may exist in agencies' program performances.

## **CONCLUSION**

### **Limitations**

The methods and designs described earlier may have limited relevance and credibility to plan for short or medium term impact assessments and impact on long-term government programs. Especially in the context of NHS Program discussed earlier. This is maybe because there are no interventions such as relevant assumptions and logical frameworks nor desired outcomes that are clearly defined or measured. In addition, government programs mostly aim to improve some aspect of complex systems, namely the economy or the environment. However, they have indeed limited control, or share responsibilities with other jurisdictions or agencies for achieving their objectives. Thus, it can be difficult to confidently attribute a causal connection between a program and its observed outcomes. In this case, government agencies need to implement a number of strategies to address evaluation challenges and develop performance information for such particular types of programs. So that, they are able to convey improved public policy implementation, better managerial aspects of efficiency and effectiveness, as well as orientation of insight and foresight for public welfare.

In respect of program evaluation methods to be more applicable, the purpose of improving them is necessary. The goal can be particularly emphasized on nuances of systematic study of the implementation and effectiveness of government policies and programs, namely a set of activities designed to achieve a specific purpose. Therefore, analysts or evaluators should bear in mind to manage challenges of evaluation design and also, among others, consider (1) strengthening the methods available to conduct evaluations, (2) promoting evidence-based policy and management, and (3) disseminating knowledge of program evaluation.

## Lessons Learned for BPK

From the previous analysis, some lessons can be learned from GAO performance audit approach, planning and design, as well as program evaluation methodology to address the questions arisen from the purpose of this paper using the topic of NHS Program such as: 1) BPK can improve its capability to conduct performance audit mainly at planning and design phase in relation to quality assurance framework by analyzing profoundly those six types of questions. This will then lead to substantiate further its approach in proposing audit objectives and current practice of developing researchable questions as regards program evaluations; 2) BPK can enhance its program evaluation methodologies using aforementioned systematic approach and research methods to collect and analyse data. This will enhance its practice in defining scope and methodology, one of which, by communicating with experts and specialists in research methods. To improve this capability, BPK could enhance its research and development capacity by hiring experts and specialists such as economists, statisticians, technologists, actuaries and legal specialists to complement its auditors or analysts; 3) BPK can increasingly add value to audited agencies and public, by designing evaluation comprehensively including the mechanism of FOD application when planning performance audit on program evaluation as regards public welfare improvement.

## REFERENCES

- Bernholz, E., et.al. (2006). *Evaluation dialogue between OMB staff and federal evaluators: Digging a bit deeper into evaluation science*. Washington, D.C. July. <http://www.fedeval.net/docs/omb2006briefing.pdf>.
- BPK. (2007). *BPK Regulation Number 01/2007 stipulating State Finance Auditing Standards*. Jakarta: BPK.
- BPK. (2009). *BPK Regulation Number 03/K/I-XIII.2/03/2009 stipulating Quality Assurance Guidance*. Jakarta: BPK.
- BPK. (2015). *BPK Regulation Number 7/K/I-XIII.2/12/2015 stipulating BPK's Strategic Plan 2016 - 2020*. Jakarta: BPK.
- BPK. (2016). *BPK Regulation Number 375/K/I-XIII.2/08/2016 stipulating Implementation Plan of the Strategic Plan 2016 - 2020*. Jakarta: BPK.
- GAO. (2008). *Policy manual*. Washington, DC: July.
- GAO. (2011a). *Government auditing standards: 2011 revision*. Washington, DC: December.
- GAO. (2011b). *Performance measurement and evaluation: definitions and relationships*. Retrieved from [www.epa.gov/sites/production/files/2015-08/documents/program-measurement-and-evaluation.pdf](http://www.epa.gov/sites/production/files/2015-08/documents/program-measurement-and-evaluation.pdf)
- GAO. (2012). *Designing evaluations 2012 revision*. Washington, DC: July.
- GAO. (2015a). *Fragmentation, overlap, and duplication: An evaluation and management guide*. Washington, DC: April.
- GAO. (2015b). *Training manual—international fellows: Developing the GAO engagement*. Washington, DC: February.
- Performance Measure. (n.d.). *In businessdictionary.com*. Retrieved 30 May 2017. <http://www.businessdictionary.com/definition/performance-measure.html>.
- Sekretariat Negara. (2003). *Act Number 17/2003 stipulating State Finance*.

Jakarta: Sekretariat Negara.

Sekretariat Negara. (2004). *Act Number 01/2004 stipulating State Treasury*. Jakarta: Sekretariat Negara.

Sekretariat Negara. (2004). *Act Number 15/2004 stipulating State Financial Accountability and Management Auditing*. Jakarta: Sekretariat Negara.

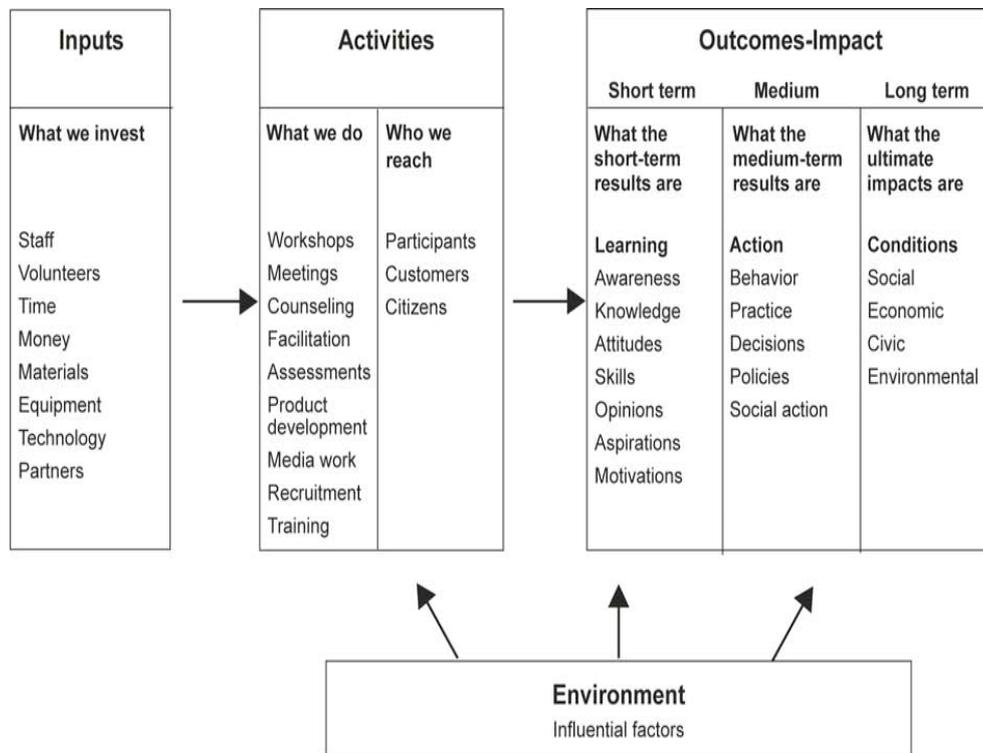
Sekretariat Negara. (2006). *Act Number 15/2006 stipulating The Audit Board of The Republik of Indonesia*. Jakarta: Sekretariat Negara.

Sekretariat Kabinet. (2015). *President Decree Number 02/2015 stipulating National Medium-Term Development Plan 2015-2019*. Jakarta: Sekretariat Kabinet.

# APPENDIX

## Appendix 1 Program Logic & Stages of Program

### Program Logic in Program Evaluation



Source: GAO and University of Wisconsin-Extension, Program Development and Evaluation.

### Common Evaluation Questions Asked at Different Stages of Program Development:

Program stage	Common evaluation questions	Type of evaluation
Early stage of program or new initiative within a program	<ul style="list-style-type: none"> <li>Is the program being delivered as intended to the targeted recipients?</li> <li>Have any feasibility or management problems emerged?</li> <li>What progress has been made in implementing changes or new provisions?</li> </ul>	Process monitoring or process evaluation
Mature, stable program with well-defined program model	<ul style="list-style-type: none"> <li>Are desired program outcomes obtained?</li> <li>What, if any, unintended side effects did the program produce?</li> <li>Do outcomes differ across program approaches, components, providers, or client subgroups?</li> </ul>	Outcome monitoring or outcome evaluation
	<ul style="list-style-type: none"> <li>Are program resources being used efficiently?</li> <li>Why is a program no longer obtaining the desired level of outcomes?</li> </ul>	Process evaluation
	<ul style="list-style-type: none"> <li>Did the program cause the desired impact?</li> <li>Is one approach more effective than another in obtaining the desired outcomes?</li> </ul>	Net impact evaluation

Source Adapted from Bernholz et al 2006.

