EVIDENCE AND EVALUATION GUIDANCE SERIES

Preparing and Appraising Evaluation Reports

A Checklist

NSW
GOVERNMENT

The NSW Ministry for Health acknowledges the traditional custodians of the lands across NSW. We acknowledge that we live and work on Aboriginal lands. We pay our respects to Elders past and present and to all Aboriginal people.

Contents

Introduction

NSW Health is committed to the generation and use of research and evaluation to improve policy and program effectiveness.* Producing evaluation reports that are high quality and relevant are key to achieving this. The _NSW Treasury Policy and Guidelines: Evaluation (TPG22-22)_[1] sets out mandatory requirements, recommendations and guidance for NSW General Government Sector agencies and other government entities to plan for and conduct the evaluation of policies, projects, regulations, and programs.

The aims of this checklist are to ensure that evaluation reports for NSW Health programs are of high quality and relevant to decision making. The checklist has two purposes:

- to inform evaluators (and those commissioned to evaluate programs) of the expectations for the preparation of evaluation reports

- to provide NSW Health staff with a framework for the appraisal of commissioned reports.

The Treasury Evaluation Policy and Guidelines outline the requirements for the evaluation of NSW public sector programs to assess their effectiveness, efficiency, value, continued relevance, and to improve transparency. The Guidelines list the recommendations for specific components that should be included in evaluation reports. This checklist complements the Population and Public Health Division's guides on _Developing and Using Program Logic_[2] and _Planning and Managing Program Evaluations_[3] for end-to-end guidance to the evaluation process.

This checklist promotes a rigorous and planned approach to the preparation and appraisal of evaluation reports. It includes criteria for ensuring that reports are complete, that the results are robust and the conclusions are sound. The checklist is adapted from the _United Nations Quality Checklist for Evaluation Reports_.[4] While it is written from the perspective of population and public health it may be useful for preparing and appraising reports from other areas such as the evaluation of clinical or health system interventions.

* In this guide the word 'program' is used interchangeably with 'initiative'. The _NSW Treasury Policy and Guidelines: Evaluation (TPG22-22)_ define an initiative as a program, policy, strategy, service, project, or any series of related events. Initiatives can vary in size and structure; from a small initiative at a single location, a series of related events delivered over a period, or whole-of-government reforms with many components delivered by different agencies or governments.

Criteria for the preparation and appraisal of evaluation reports

The following checklist contains criteria to guide the preparation and appraisal of evaluation reports. The range of criteria is extensive and not all reports will need to be appraised against all of the criteria. For example, when appraising a final report of a complex evaluation where the methods are described in previous or complementary reports, you may not expect the methods to be reported in the same level of detail in the final report. This checklist can be used as a guide, taking into consideration the other contextual factors of the evaluation. Before a report is written it is important for all stakeholders to understand the requirements of the report and the intended audience.

The criteria are arranged under nine headings that correspond to the contents of a typical evaluation report submitted to NSW Health.

To save your inputs, **download and save a copy** of the PDF before you begin. Information entered directly into a web browser will not save.

How to use the checklist
Planning and commissioning reports

This checklist can be used in the planning and/or commissioning stage of an evaluation. For example:

- include the checklist in the planning and commissioning documents so evaluators are clear on expectations for the final report

- use the checklist as a discussion tool with an evaluator when deciding what should and should not be included (defining the requirements)

- you may make it a requirement that evaluators submit the completed checklist with the final report.

The checklist can also support performance management of evaluators throughout the evaluation process. The checklist provides clear expectations of what is required and can be used to guide conversations when these are not being met.

Writing reports

The checklist can be used to ensure you address the criteria for a high quality report during the writing stage. Go through the checklist as you prepare the report to make sure you meet these standards.

Appraising reports

Review evaluation reports against the criteria in the checklist to ensure they are complete and high quality. It may not be necessary for a report to meet all criteria depending on the type and purpose of the evaluation; this should be considered and agreed upon during the evaluation planning stage.

- Mark if each criteria is addressed in the report: 'yes', 'no', 'insufficient' or 'not applicable (N/A)'.

- There is space for recording comments to support your assessment if needed.

- If required criteria are not met, use the checklist and your comments to guide a discussion with the evaluator to address the issue.

A note on report format

Reports should be as concise as possible. One way to achieve this is to follow a 1:3:25 pages plus appendices format. This format includes:

- a one page summary of key implications

- a two to three page executive summary

- 25 pages to cover the introduction, evaluation methods, results, conclusions, and recommendations (if requested) with supporting documents provided as appendicies.

Clear and succinct writing, and use of appendices for supporting material can help an evaluator keep the body of a report to 25 pages. For more guidance on the 1:3:25 format, see the *Reader-Friendly Writing - 1:3:25 tool*.

Please note, there may be instances where 25 pages for the body of the report is not suitable and more or less information is required. For example, evaluations of large-scale multi-component programs may be best presented as a separate report for each component. Consider the needs of each evaluation and communicate expectations with the evaluator early in the process.

Evaluation report checklist

Program:	
Agency:	
Evaluator:	

1. Report structure		
1.0	Report is well structured, logical, clear and complete	
1.1	Report is logically structured with clarity and coherence (e.g. background and objectives are presented before results, and results are presented before conclusions and recommendations).	
1.2	Title page and opening pages provide key basic information, including: • name of the program†, evaluators and commissioning agency • timeframe of the evaluation and date of the report • table of contents, list of acronyms and glossary.	
1.3	One-page outline includes the key lessons decision makers can take from the evaluation.	
1.4	Executive Summary is a stand-alone section of 2-3 pages, including: • overview of the program • evaluation objectives and intended audience • brief description of the evaluation methods • most important results and conclusions • main implications for practice (if requested).	
1.5	Body of report is approximately 25 pages (if appropriate).	
Comments:		

2. Program evaluated		
2.0	Report presents a clear and full description of the evaluated program	
2.1	Brief introduction/background to the topic summarising current knowledge on the topic (no more than half a page).	
2.2	Key social, political, economic, demographic and institutional factors that have a direct bearing on the program are described. For example: • relevant government and/or commissioning agency strategies and priorities • acknowledgement of previous evaluation reports and a short summary of findings.	
2.3	Program logic model is included and described. See *Developing and Using Program Logic: A Guide*[2] for further information on program logic models.	

---

†In this checklist, the term 'program' will be used to refer to both policies and programs.

2.4	Clear description of the program being evaluated. This should include the scale and complexity of the program. For example:	
	• purpose, goal, and organisation/management of the program	
	• description of all components of the program, including delivery methods and intervention dose	
	• size and characteristics of the population each component is intended to serve, either directly or indirectly	
	• setting or geographic context such as the LHD(s) or region(s) where the program was implemented, and challenges where relevant	
	• total resources from all sources, including human resources (e.g. number of service providers) and budget(s) (including Agency, State and Commonwealth funding if relevant).	
2.5	Key program stakeholders and their roles (e.g. implementing agencies and/or partners).	
2.6	Description of the implementation status of the program, for example:	
	• phase of implementation	
	• any significant changes (e.g. to plans, strategies and/or logical frameworks)	
	• implications of the changes for the evaluation.	
Comments:		

3. Evaluation purpose and scope

3.0	Evaluation's purpose, objectives and scope are fully explained	
3.1	Purpose of the evaluation is clearly defined including:	
	• why the evaluation was needed at that point in time	
	• who needed the information	
	• what information was required	
	• how the information is intended to be used.	
3.2	A clear explanation of the evaluation objectives, scope, evaluation criteria and key evaluation questions. The report describes and justifies what the evaluation did and did not cover.	
3.3	Management and governance of the evaluation is described.	
Comments:		

4. Evaluation methods

4.0	Report presents a transparent description of the methods that clearly explains how the evaluation was designed to address the evaluation criteria, answer the evaluation questions and achieve the evaluation purpose.\n\nSee 'Reporting Guidelines' section for more information.	
4.1	Study design is specified at the beginning of the methods section (e.g. randomised control trial, mixed methods).\n\nSee *Study Design for Evaluating Population Health and Health Service Interventions: A Guide*[5] for more information	
4.2	Primary and secondary outcomes or the measures of effect used to answer the key evaluation questions are described.	
4.3	Details about data collection are described, including:	
	• data collection methods	
	• data analysis methods (i.e. the specific quantitative and/or qualitative analysis approaches, software used)	
	• rationale and limitations for chosen methods	
	• baseline data and performance indicators (where relevant).	

4.4	Data sources, the rationale for their selection and their limitations are described. The report includes a discussion of how the mixture of data sources was used to: • obtain diverse perspectives on the program • ensure data accuracy • address data limitations.	
4.5	Approach to sampling for the evaluation is described, including: • area(s) and population(s) to be represented • rationale for the selection of the sample(s) • sampling methods used • numbers selected out of potential subjects • limitations of the sample(s).	
4.6	Process of involving stakeholders in evaluation planning, implementation and interpretation of results is described.	
4.7	Methods are appropriate for the evaluation purpose and to answer the evaluation questions.	
4.8	Demonstrate that research methods were applied rigorously. If there are limitations to the rigour of the research methods, these are reported and the implications are discussed (see checklist 5.6). See 'Critical Appraisal Tools' section for more information and suggested resources.	
4.9	Evidence is presented that adequate measures were taken to ensure data quality, including ensuring the reliability and validity of data collection tools (e.g. interview protocols, observation tools, approaches to cleaning and preparing data).	
4.10	Ethics approvals and other institutional approvals that were obtained for the evaluation are listed (where applicable).	
4.11	There is a statement addressing any potential for a conflict of interest in the conduct of the evaluation.	

Comments:

5. Results		
5.0	Results correspond directly to the evaluation criteria and questions detailed in the scope and objectives section of the report. The results are based on evidence derived from data collection and analysis described in the methods section of the report.	
5.1	Results reflect a systematic and appropriate analysis and interpretation of the data	
5.2	Results are reported for all planned analyses (as per methods section).	
5.3	Reporting of quantitative results includes: • statistical tests are reported and interpreted correctly, including significance tests (i.e. p-values, confidence intervals) • effect sizes are reported.	
5.4	Reporting of qualitative findings includes: • data is analysed to identify themes and sub-themes • findings are reported with a statement that identifies it as a general theme or applying to just one respondent • quotes are included to illustrate key findings. Quotes are de-identified but include signifier e.g. [participant 4] or [4, Male, age 45]) • findings are triangulated against other sources of data (these may be quantitative and/or qualitative).	
5.5	Reporting of economic appraisal information includes: • approaches for defining, calculating and costing all program costs and benefits are clearly specified • if health benefits have been monetised (i.e. for cost-benefit analysis), the method and results of conversion (e.g. value of a statistical life year, willingness to pay studies) are clearly specified with a rationale for appropriateness • metrics such as incremental cost-effectiveness ratios (ICERs) or benefit-cost ratios (BCRs) are interpreted for the reader, so someone without any background in health economics can understand the implications of the economic appraisal.	

5.6	Results are objectively reported based on the evidence and includes the extent that results can be attributed to the program.	
5.7	Gaps and limitations in the data and/or unanticipated results are reported and discussed.	
5.8	Reasons for accomplishments and challenges/failures, especially continuing constraints, are identified as much as possible.	
5.9	Overall, results are presented with clarity, logic and coherence, and address the evaluation questions.	
Comments:		

6. Conclusions

6.0	Conclusions present reasonable judgements based on results and substantiated by evidence, and provide insights pertinent to the program and purpose of the evaluation.	
6.1	Conclusions provide clear and reasoned insights that relate to the evaluation questions.	
6.2	Conclusions are well substantiated by the evidence presented and are logically connected to evaluation results.	
6.3	Conclusions give insights into issues that are relevant to the evaluation users.	
6.4	Conclusions present strengths and weaknesses of the program being evaluated, based on the evidence presented and taking due account of the views of a diverse cross-section of stakeholders.	
Comments:		

7. Implications for practice[†] (if requested)

7.0	Implications for practice are relevant to the program and purposes of the evaluation, are supported by evidence and conclusions, and were developed with the involvement of stakeholders.	
7.1	The report describes the process followed in developing the implications for practice including consultation with stakeholders.	
7.2	Implications for practice are firmly based on evidence and conclusions.	
7.3	Implications for practice are relevant to the program and purpose of the evaluation.	
7.4	Implications for practice are actionable and reflect an understanding of the context.	
Comments:		

8. References

8.1	References for all information in the report that originally appeared in other sources.	
8.2	References are presented in a recognised standard format (preferably Vancouver format)	
Comments:		

[†]'Implications for practice' can be used in the place of 'Recommendations'. This section considers the results and conclusions of the evaluation in the context of the program and potential future activity. They are less directive than recommendations.

9. Appendices	
9.1	Appendices increase the credibility of the evaluation report.
	Appendicies may include (but are not limited to):
	• Program logic model
	• Evaluation plan/data matrix
	• Sampling and recruitment approach
	• More details of the methodology, such as data collection instruments (or citations if these are lengthy), including details of reliability and validity
	• Evaluation terms of reference
	• Governance structures
	• Additional data tables
	• List of respondents interviewed and sites visited (only if appropriate)
	• List of documents consulted or reviewed
Comments:	

Further resources

Formatting and terminology

Planning and commissioning reports

More information about how to structure a report using the 1:3:25 approach and the benefits of this approach is available at:

- *Canadian Health Services Research Foundation – Reader-Friendly writing – 1:3:25*

For guidance on appropriate terminology to use for a culturally safe report in regards to Aboriginal patients, staff and communities please review the following guide:

- *Communicating Positively: A Guide to Appropriate Aboriginal Terminology*

Reporting guidelines

Reporting guidelines help both the writer and reader to make sure all the relevant information about an intervention and study are included in a report.

Interventions

There are specific intervention features that should be reported when describing population health and policy interventions to ensure maximum transparency and replicability. The TIDieR-PHP checklist outlines these details.

Study designs

Different study designs require different information to be reported about the methods and results. If you are unsure what should be reported for a particular study design, the following tools may be a good reference point.

- Randomised control trials – CONSORT

- Quasi-experimental studies (non-randomised designs) – TREND

- Cohort, case-control or cross-sectional studies – STROBE

- Economic evaluations – CHEERS

- Qualitative studies – SRQR (general qualitative designs); COREQ (interviews and focus groups)

- Further study designs are available at the Equator Network

These reporting guidelines are designed for academic manuscripts so some sections, in particular guidance relating to the discussion of findings, may not be relevant to appraising evaluation reports. When using these tools with the checklist in this guide, focus on the 'methods' and 'results' sections.

Critical appraisal tools

Checklist item 4.8 asks us to appraise how rigorously the methods have been applied. This is important as an indicator of the validity of the results and the quality of the evidence generated. Critical appraisal tools can help us apply a systematic approach to making these judgements.

If you are unsure of what to consider when making this appraisal, there are a number of critical appraisal tools available online (listed below). You can also consult with a research and/or evaluation expert in your agency for guidance.

- JBI checklists
- CASP checklists
- National Heart, Lung and Blood Institute assessment tools

References

1. NSW Treasury. *NSW Government Policy and Guidelines: Evaluation*. Sydney: NSW Treasury; 2023. Available from: www.treasury.nsw.gov.au/finance-resource/evaluation-policy-and-guidelines

2. Centre for Epidemiology and Evidence. *Developing and Using Program Logic: A Guide*. Evidence and Evaluation Guidance Series, Population and Public Health Division. Sydney: NSW Ministry of Health; 2023. Available from: www.health.nsw.gov.au/research/Pages/developing-program-logic.aspx

3. Centre for Epidemiology and Evidence. *Planning and Managing Program Evaluations: A Guide*. Evidence and Evaluation Guidance Series, Population and Public Health Division. Sydney: NSW Ministry of Health; 2023. Available from: www.health.nsw.gov.au/research/Pages/planning-evaluations.aspx

4. United Nations Evaluation Group. *United Nations Quality Checklist for Evaluation Reports*. United Nations, Geneva; 2010. Available from: www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/about_iom/eva_techref/UNEG_Eval_Report.pdf

5. Centre for Epidemiology and Evidence. *Study Design for Evaluating Population Health and Health Service Interventions: A Guide*. Evidence and Evaluation Guidance Series, Population and Public Health Division. Sydney: NSW Ministry of Health; 2023. Available from: www.health.nsw.gov.au/research/Pages/study-design-guide.aspx