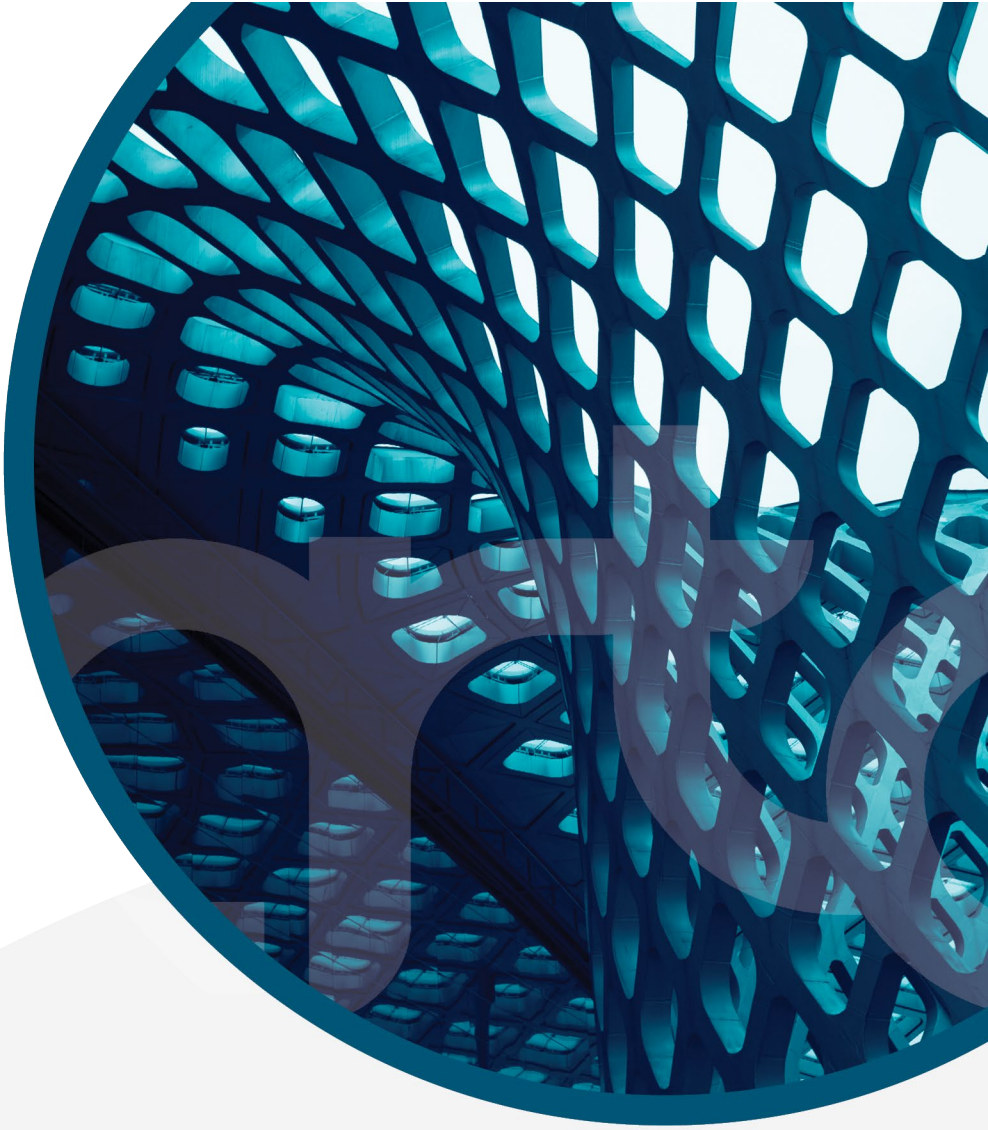


NATIONAL MENTAL HEALTH AND SUICIDE PREVENTION EVALUATION FRAMEWORK

FINAL	DECEMBER 2023

www.artd.com.au

CONTENTS

1. Framework on a page 3

2. Background 4
   - 2.1 National Mental Health and Suicide Prevention Agreement 4
   - 2.2 A National Evaluation Framework and Sharing Guidelines 4
   - 2.3 How the Framework was developed 6

3. Purpose and principles for evaluation 7
   - 3.1 What evaluation is 7
   - 3.2 How evaluation fits with other evidence generating activities 8
   - 3.3 A principles-based approach to evaluation of mental health and suicide prevention programs 10

4. Getting ready for evaluation 13
   - 4.1 Ensuring readiness for evaluation 13
   - 4.2 Deciding whether and when to evaluate a program 18
   - 4.3 Establishing governance arrangements 19

5. Deciding the focus of evaluation 22
   - 5.1 Focus areas for evaluation 22
   - 5.2 Key evaluation questions 26

6. Designing a fit-for-purpose evaluation 29
   - 6.1 Considerations to inform evaluation approaches and designs 29
   - 6.2 Evaluation approaches and designs 31

7. Choosing appropriate methods and measures 45
   - 7.1 Evaluation methods 45
   - 7.2 Identifying appropriate measures 47
   - 7.3 Engaging service users in evaluation 68

8. Ensuring the evaluation is conducted ethically 77
   - 8.1 Key ethical considerations 77
   - 8.2 Formal ethics approval 78
   - 8.3 Ethical requirements for evaluations involving Aboriginal and Torres Strait Islander communities 79

9. Analysing and reporting 81
   - 9.1 Analysis 81
   - 9.2 Keeping stakeholders informed of progress 83
   - 9.3 Reporting 83

Appendix 1 Key concepts 84

Appendix 2 Other relevant documents 86

Appendix 3 Engaging lived and/or living experience team members 87

Appendix 4 Evaluability assessment checklist 95

Appendix 5 Outcomes mapping 97

Appendix 6 EMMIE template 103

1. FRAMEWORK ON A PAGE

PURPOSE

To help improve the mental health and suicide prevention system for the people and communities who need them. It will do this by strengthening the quality and consistency of evaluations of mental health and suicide prevention initiatives.

EVALUATION PRINCIPLES

- Lived and living experience and consumer centred
- Robust and fit for purpose
- Trauma responsive and ethical
- Focused on learning and building the evidence base

GETTING READY FOR EVALUATION

It is important to consider evaluation upfront to ensure that appropriate data is collected to understand how a program is working and whether outcomes are achieved. It is also important to build in appropriate governance that includes various levels of government involved, lived and living experience and Aboriginal and Torres Strait Islander governance.

DOMAINS OF FOCUS FOR EVALUATION

Areas of focus for evaluation to provide a common language for comparison so that evaluations – big and small – can contribute to our understanding of how to improve activities, initiatives and the mental health and suicide prevention service system.

EVALUATION APPROACHES AND DESIGNS

Advice on choosing an evaluation approach and design that is fit for purpose - that suits the scale of the program, context, stage of development and other considerations.

ETHICAL CONSIDERATIONS

Considerations to support ethical evaluation, trauma-informed approaches with people with lived experience, and culturally safe evaluation for Aboriginal and Torres Strait Islander communities.

METHODS AND MEASURES

Guidance on making use of commonly available data for mental health and suicide prevention initiatives, and additional data collection evaluations might consider.

ANALYSIS AND REPORTING

Guidance on analysis and interpretation, reporting of evaluations to help build the evidence base.

2. BACKGROUND

2.1 NATIONAL MENTAL HEALTH AND SUICIDE PREVENTION AGREEMENT

The National Mental Health and Suicide Prevention Agreement (National Agreement) sets out the shared intention of the Commonwealth, state and territory governments to work together to:

- improve the mental health of all Australians
- reduce the rate of suicide towards zero
- improve the Australian mental health and suicide prevention system.

The activities implemented as part of the National Agreement will be **informed by people with lived and/or living experience** of mental ill-health and/or suicide, and their families, carers and support people.

Under the National Agreement, governments have committed to measure and report on their activities and the outcomes of these.

2.2 A NATIONAL EVALUATION FRAMEWORK AND SHARING GUIDELINES

Under the National Agreement, governments agreed to develop a National Evaluation Framework (the Framework) and National Evaluation Sharing Guidelines (the Guidelines) (Annex C). The Framework relates to a key recommendation from the *Productivity Commission 2020 Inquiry into Mental Health* to embed an evaluation culture into the system.

2.2.1 PURPOSE OF THE FRAMEWORK

The purpose of this Framework is to help improve the mental health and suicide prevention system for the people and communities who need it. The Framework will do this by providing guidance to strengthen the quality and consistency of evaluations of mental health and suicide prevention programs.

While the Framework communicates what is expected of evaluations under the National Agreement, it is intended to be useful beyond the timeframe of the National Agreement to guide the evaluation of any mental health and suicide prevention program. The Framework is not intended to be prescriptive but to guide those planning, commissioning, and conducting evaluations.

Australian jurisdictions are at various stages of implementing mental health and suicide prevention programs and systematic reform. This means that each jurisdiction will apply the Framework differently. The Framework is designed to be used alongside the program implementation and systematic reforms underway within jurisdictions.

The Framework is intended to:

- support more **consistent** evaluation of mental health and suicide prevention programs, including more consistent approaches to measuring effectiveness and efficiency, which in turn will build a stronger evidence base and support meta-analysis where useful
- **reduce the burden** of data collection for organisations, particularly small organisations, receiving funding from multiple sources by encouraging consistency
- help to strengthen the **quality** of evaluations of mental health and suicide prevention programs by building a shared understanding of quality evaluations of these programs and supporting evaluation capability building
- through consistency of data collection, enable a better view of the evidence at the **system-level** by enabling evaluations of diverse programs to be compared and synthesised
- encourage the **use** of evaluations for improvement and decision-making
- support the role of people with **lived and/or living experience** of mental ill-health and suicide, including carers and support people of those with lived and/or living experience in the design, conduct and reporting of evaluation through guidance about a trauma-informed approach to this.

The Framework is not a framework for reporting on the evaluation of the National Agreement, as this is a separate piece of work.

For programs funded under the National Agreement, the roles and responsibilities in evaluation are outlined in the bilateral agreements. The National Agreement suggests that consideration be given to cost sharing of evaluations of programs of national significance.

2.2.2 INTENDED USERS

The Framework is intended to be useful to all groups with a stake in mental health and suicide prevention:

- **government agencies** implementing programs under the National Agreement to guide planning and commissioning of evaluations
- government funded or delivered **mental health and suicide prevention** service providers to guide data collection, planning, and commissioning of their own evaluations, and participation in government-commissioned evaluations
- **researchers and evaluators** to design and deliver evaluations more consistently
- **people with lived and/or living experience** of mental ill-health and/or suicide, and their families, carers, and support people, to understand how mental health and suicide prevention programs may be evaluated, and how they can contribute to or participate in the evaluation
- **other agencies** to consider incorporating mental health outcomes when evaluating their programs, given the important impact that things like housing and employment supports can have on mental health.

2.3 HOW THE FRAMEWORK WAS DEVELOPED

Mental Health and Suicide Prevention Senior Officials (MHSPSO) established an Evaluation Project Group to guide the development of the National Evaluation Framework.

ARTD Consultants – with academic partners (Myfanwy Maple and Sarah Wayland), linked data specialists (Taylor Fry), an Indigenous Consultant (Tom Brideson), and a network of lived and/or living experience team members – was selected by the Australian Government Department of Health and Aged Care to draft the Framework.

The team reviewed existing evaluations and evaluation guidance from the Australian Government and state and territory governments, and consulted with stakeholders to ensure the Framework builds on existing good practice and helps to overcome current concerns with evaluation. Key stakeholder groups consulted included:

- representatives from the Australian Government, state and territory governments
- Primary Health Networks
- lived and/or living experience and priority population representative groups
- state, territory and national Mental Health Commissions
- state, territory and national peak bodies and mental health and suicide prevention organisations
- researchers and evaluators
- mental health and suicide prevention data custodians.

3. PURPOSE AND PRINCIPLES FOR EVALUATION

All government funded mental health and suicide prevention programs should be able to demonstrate they are delivering evidence-based interventions safely, effectively, and efficiently, to the people and communities who need them[1]. Evaluation – alongside other evidence generating activities – can help to demonstrate what is working well and what needs to be improved. It is essential for understanding the effectiveness of existing programs and systems, and for supporting development and comparison of new service models.

This section outlines the purpose evaluations can be used for, and how evaluation can complement other methods for generating evidence about how programs are working.

3.1 WHAT EVALUATION IS

Evaluation is an in-depth process for determining the merit, worth or value of a program, policy or strategy[2]. It involves systematic collection of information about a program to answer key evaluation questions.

3.1.1 TIMING OF EVALUATION

Often people think of evaluation as something that is done at the end of a program, but evaluation is useful across the lifecycle of a program, from early planning and development through implementation and decision-making about ongoing funding or scale-up.

3.1.2 PURPOSE OF EVALUATION

It can be used to:

- help strengthen the **design of a program** – for example, through a needs analysis to ensure the program is addressing the needs of the target group, the development of a logic model that shows how the program is expected to contribute to its intended outcomes, and identification of the existing evidence for a program model
- help strengthen **implementation of a program** – by identifying what is working well, what is not, and how the program can be improved, including opportunities to better reach priority populations and improve the efficiency of the program
- help inform **decision-making about ongoing funding** – by identifying whether it is achieving the intended outcomes, and the value for money provided by the program
- help inform **considerations about scale-up** – by identifying the critical elements of a program and the contextual factors important for implementation to succeed
- enable **accountability to funders and people accessing programs** – by identifying what they are achieving
- accountability to key stakeholders and the wider public.

[1] Federal Financial Relations (2022). National Mental Health and Suicide Prevention Agreement. https://federalfinancialrelations.gov.au/agreements/mental-health-suicide-prevention-agreement
[2] Scriven, M. (1991). *Evaluation thesaurus*, 4th ed. USA: Sage Publications, Inc.

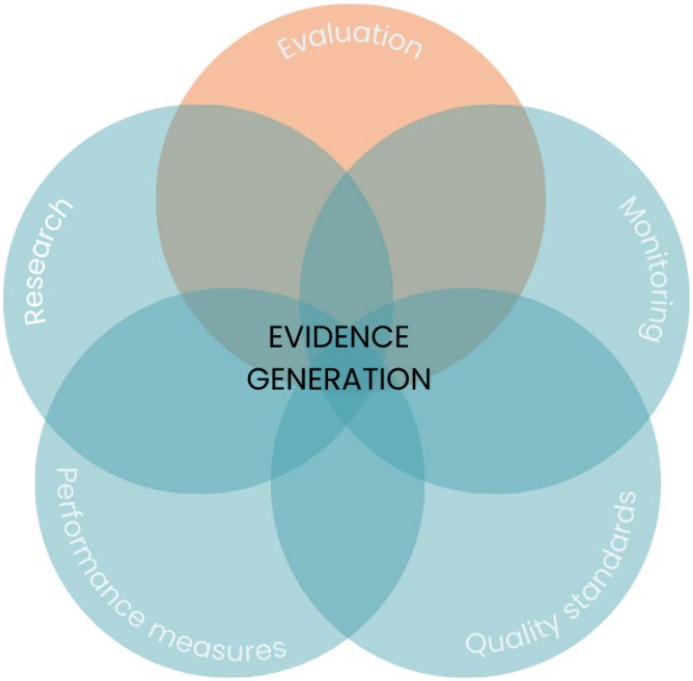
When shared, evaluations can help to build the **knowledge base** for the mental health and suicide prevention system, identifying what is working, for whom and in what circumstances.

3.2 HOW EVALUATION FITS WITH OTHER EVIDENCE GENERATING ACTIVITIES

In addition to evaluation, there are various evidence-generating activities that help governments and organisations improve their practice and remain accountable. While they have different specific purposes, many of these share some underlying data collection techniques with evaluation. Other forms of evidence can also be an important source of information for evaluation.

Figure 1 highlights how evaluation fits in with other evidence-generating activities.

FIGURE 1. FIT BETWEEN EVALUATION AND OTHER EVIDENCE GATHERING ACTIVITIES



Source: Adapted from Department of Finance (2021). *Evaluation in the Commonwealth (RMG 130).*

3.2.1 MONITORING

Monitoring is the routine collection of program performance information during implementation to determine whether a program is on track and whether any adjustments are required.[3] Monitoring:

[3] Department of Health and Aged Care (2022). Evaluation Strategy 2023-26.

- mainly uses data captured as part of delivering the program
- can inform quality assurance activities
- should be used to manage activities on an ongoing basis
- can contribute to continuous improvement
- can help to assess whether a program is on track to achieve intended outcomes.

In mental health and suicide prevention, there are a range of minimum data sets collected by the Australian Government, state and territory governments and funded services.[4] Several data collections contain outcomes and case mix measures, which provide information about an individual consumer based on their engagement with a mental health service. For example, the *Mental Health National Outcomes and Casemix Collection* **(NOCC)**, and the *Primary Mental Health Care Minimum Data Set* **(PMHC MDS)** (see section 7.2.3). There is also the National Suicide and Self-Harm Monitoring System established by the National Suicide Prevention Office, together with the Australian Institute of Health and Welfare.

Monitoring data is an important source of data for evaluation but is not in itself evaluation. Monitoring data may indicate the need for an evaluation to provide greater insight into why the program or parts of the program are producing different results to what was expected. Monitoring data also feeds into system performance measurement through the use of indicators collected as part of minimum data sets.

3.2.2 PERFORMANCE MEASURES

Performance measures generally make use of a smaller subset of monitoring data. At the national level, the *Health Performance Framework* and the *National Mental Health Performance Framework 2020* define domains of routinely collected data in health and mental health.

Specific performance measures can also be developed for individual services.

3.2.3 RESEARCH

Research can be an important input to evaluation to:

- assess the appropriateness of a program design
- identify existing data collection tools to use in evaluation
- compare the outcomes of a program to other programs.

Evaluations can also raise questions that research would be best placed to answer, for example, about best practice or long-term outcomes. In mental health and suicide prevention, national research and clinical trials are funded through national research bodies, such as the National Health and Medical Research Council, the Medical Research Future Fund, and the Australian Research Council. Post evaluation research may include qualitative, quantitative, and/or mixed methods research.

[4] Federal Financial Relations (2022). National Mental Health and Suicide Prevention Agreement. https://federalfinancialrelations.gov.au/agreements/mental-health-suicide-prevention-agreement

3.2.4 QUALITY STANDARDS

Information produced for reporting against **quality standards** can feed into an evaluation of the implementation of a program.

In mental health, the *Australian Commission on Safety and Quality in Health Care* sets the *National Standards for Mental Health Services* in health settings. The *National Safety and Quality Health Service (NSQHS) Standards* provide a nationally consistent statement of the level of care consumers can expect from health service organisations. *The National Safety and Quality Digital Mental Health (NSQDMH) Standards* aim to improve the quality of digital mental health service provision, and to protect service users and their support people from harm.

In suicide prevention, Suicide Prevention Australia (a national peak body) also has a voluntary, independent quality accreditation scheme. This scheme is intended to support organisations to implement safe, high-quality, and effective suicide prevention and postvention programs in Australia.

3.3 A PRINCIPLES-BASED APPROACH TO EVALUATION OF MENTAL HEALTH AND SUICIDE PREVENTION PROGRAMS

It is not possible to define a standard approach to evaluation for the range of mental health and suicide prevention programs because of the diversity of programs and their intended outcomes, the diversity of purposes for evaluation, and the diversity of contexts for evaluation. Instead, this Framework outlines some principles to guide evaluation and considerations for each stage of evaluation.

3.3.1 PRINCIPLES TO GUIDE EVALUATION

The following principles are suggested to guide decision-making and support the evaluation of mental health and suicide prevention programs:

- **Lived and/or living experience and person-centred**
    - People with lived and/or living experience of mental ill-health and/or suicide are actively involved in the design and delivery of evaluations to ensure they are useful and meaningful. (This may include co-design or co-production).
    - Evaluation processes are safe and accessible to people accessing the service, including for diverse populations.
    - Evaluations include families, carers and supporters, where relevant.
- **Robust and fit-for-purpose**
    - Evaluation is considered upfront when a service is being designed to ensure that data collection is built in, and outcomes measures are meaningful.
    - Evaluations make the best use of existing data to minimise the burden on people accessing the service and the workforce.
    - Evaluations are designed to suit the stage of development of the program, questions to be answered for decision-making, available budget, timeframe, and context.

  - Evaluations are designed with context in mind, and findings are considered with reference to context.
- **Trauma-responsive and ethical**
  - Evaluations are designed in a way to avoid traumatisation and/or re-traumatisation.
  - The evaluation follows ethical standards, including those outlined in National Health and Medical Research Council guidelines[5] and specific guidance for Aboriginal and Torres Strait Islander communities.
  - The benefits of participating in the evaluation exceed the potential risk of harm.
- **Focused on learning and building the evidence base**
  - Evaluations enable service improvement by being clear about what is working well and what needs to be improved and the context in which this is occurring.
  - Evaluations are conducted in a way that they can contribute to the broader evidence base about mental health and suicide prevention programs in Australia.
  - Evaluations provide clear and actionable implications or recommendations that can be shared with relevant stakeholders.

3.3.2 STEP BY STEP CONSIDERATIONS FOR EVALUATION

The rest of this Framework outlines key considerations for making decisions about evaluation that are fit-for-purpose.

[5] The National Health and Medical Research Council, Australian Research Council and Universities Australia (2023). National Statement on Ethical Conduct in Human Research. Canberra: National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2023

GETTING READY FOR EVALUATION

- Getting ready for evaluation helps ensure you have the data needed to understand whether a program is working
- From the outset, identify intended outcomes and the data required to understand implementation and outcomes
- Confirm whether and when to evaluate based on level of funding, risk and available evidence
- Establish appropriate governance, where multiple governments are involved, people with lived and living experience and Aboriginal and Torres Strait Islander communities

DEFINING THE FOCUS OF EVALUATIONS

- Define the focus for evaluation – design, process, outcomes and/or economic – based on timing and information needs
- Identify the key evaluation questions that address the needs of intended users and can inform decision-making
- Some consistency across evaluations will support a system-level view

DESIGNING A FIT- FOR-PURPOSE EVALUATION

- Identify an approach that is appropriate for the nature and scale of the program, context, questions to be answered and time and resources
- Consider the most appropriate approach to assessing outcomes, given the context
- Mix methods if possible and appropriate

CHOOSING APPROPRIATE METHODS AND MEASURES

- Make the best use of available data on processes and outcomes and costs
- Supplement existing data with new data collection to fill gaps
- Take a trauma-informed approach to collecting data from service users

ENSURING THE EVALUATION IS CONDUCTED ETHICALLY

- Align with NHMRC guidance, ensuring informed consent, voluntary participation and doing no harm
- Ensure specific ethical considerations when working with Aboriginal and Torres Strait Islander communities, including Indigenous data sovereignty

ANALYSING AND REPORTING

- Synthesise all evidence sources
- Consider how actual implementation and outcomes compare to what was intended
- Involve stakeholders in interpretation of findings and their implications
- Clearly communicate findings in formats suitable for the audience

4. GETTING READY FOR EVALUATION

This chapter helps to identify whether and when to evaluate a program, to collect the data required to understand how a program is working whether or not it is evaluated, and to prepare for evaluation if you decide an evaluation is required. This should be read in conjunction with any jurisdiction-specific evaluation guidelines, as well as funding contract requirements around monitoring and evaluation.

Useful references

Different jurisdictions have their own guidance on conducting evaluations.

- *ACT Government: Evaluation Policy and Guidelines*
- *Evaluation in the Commonwealth*
- *Northern Territory Government: Program Evaluation Framework*
- *NSW Treasury Evaluation Policy and Guidelines*
- *Queensland Government Program Evaluation Guidelines*
- *South Australian Government Economic Insight and Evaluation*
- *Tasmanian Government: Evaluation resources*
- *Victorian Evaluation resources*
- *Government of Western Australia: Program evaluation*

4.1 ENSURING READINESS FOR EVALUATION

Too often, evaluations report that data was of insufficient quality to draw conclusions. To get the most out of an evaluation – and to ensure that there is the data required to understand whether a program is working, regardless of whether an evaluation is undertaken – organisations should plan and build in quality data collection from the outset.

4.1.1 THE BENEFITS OF BUILDING IN CONSIDERATION OF EVALUATION EARLY

When evaluation is considered during the development of a program, it can:

- clarify the need for a program
- help to clarify how the program is intended to operate or strengthen the program's design
- clarify the intended outcomes of a program and how these can be measured
- ensure that useful baseline and monitoring data are identified and collected, with the right permissions to enable use in evaluation
- help to ensure the information needs of stakeholders can be met
- ensure that adequate resources are available to undertake high-quality evaluations.

While evaluations can be considered at a later stage, this may affect the quality of data available, and limit potential evaluation designs, making it harder to answer key evaluation questions.

4.1.2 ASSESSING READINESS FOR EVALUATION

Knowing the quality of monitoring data (see section 3.2.1) might be enough to proceed with an evaluation of a small-scale program.

For larger-scale programs, to ensure you get value out of an evaluation, you might consider a more formal evaluability assessment: "an assessment of the extent to which an intervention can be evaluated in a reliable and credible fashion."[6]

There are different approaches to conducting an evaluability assessment, but broadly they include:

- whether the program can be evaluated given the nature of the design
- whether there is relevant data and the data collection systems to provide data for evaluation
- whether it is practical and useful to conduct an evaluation given stakeholder views and availability.

See Appendix 4 for an example of an evaluability assessment checklist. The following sections outline key steps to ensure readiness for evaluation and establish useful data about a program, whether or not an evaluation is undertaken.

4.1.3 IDENTIFY KEY STAKEHOLDERS

In preparing for an evaluation, it is important to identify the key stakeholders:

- the intended users of the evaluation (such as funders, program managers and staff)
- those who will contribute to the evaluation (such as service users, staff, and broader stakeholders, including those referring to or coordinating with the program)
- those who will have an interest in the evaluation.

This can help to identify key information needs for the evaluation and ensure the relevance and credibility of the evaluation from stakeholders' perspectives.

Intended users of the evaluation will have an important role in deciding the focus of evaluation (see section 5). Engaging intended users of an evaluation upfront can help to enhance use of the evaluation.

4.1.4 IDENTIFY INTENDED OUTCOMES

It is important for programs to be clear about their intended outcomes – that is, what they intend to change – from the outset.

Developing a **logic model** can help to clarify what outcomes a program is intended to achieve. A logic model is a one-page diagram that summarises how a program is intended to

[6] OECD-DAC (2010). Glossary of Key Terms in Evaluation and Results Based Management. Paris: OECD-DAC. https://www.oecd.org/development/peer-reviews/2754804.pdf

work. It describes the inputs, the activities, and the outputs that the program intends to deliver, and the sequence and links between the outcomes (for the people accessing the program, the workforce, and society). Figure 2 provides an example of a program logic model – although there are many ways to represent a program on a page.

Some government agencies require that a logic model is developed as part of initial funding proposals or program development because it can help to strengthen a program's design and likelihood of success.

Collaboratively developing a logic model with key stakeholders, including people intended to benefit from the program and staff, can help to ensure the program design is feasible and outcomes are meaningful.

FIGURE 2. PROGRAM LOGIC MODEL EXAMPLE

Source: Adapted from Department of Finance (2021). *Evaluation in the Commonwealth (RMG 130).*

Useful references

- The jurisdiction-based evaluation guidelines referenced above
- *Creating program logic models | Better Evaluation*

4.1.5 ESTABLISH THE DATA REQUIRED AND MAKE USE OF IT

Developing a logic model can help to identify what data will be required to assess whether the program is being implemented as intended, and whether the intended outcomes are being achieved (i.e. monitoring data). This involves identifying data that will be used to assess each 'box' in the logic model.

Sections 7.2.2 and 7.2.3 outline standard process and outcomes measures for mental health and suicide prevention programs that can be drawn on in establishing important data about a program.

Considering these data requirements upfront provides an opportunity to ensure data requirements are included in funding agreements with service providers and built into data systems. It can also help to ensure data is collected with the right permissions, so that it can be used for evaluation and service users can be followed up for evaluation.

Considering data requirements upfront can also provide the opportunity for co-creation of measures so that they are acceptable to people who access programs and to staff, are feasible to implement, and are valued.

Regularly reviewing monitoring data as part of implementation can help to identify whether the program is on track or if adjustments are needed. It can also help to ensure that data is being collected as intended and is of good quality for use in evaluation – for example, that there aren't major gaps, and that questions are being interpreted consistently.

4.1.6 CONFIRM AVAILABLE RESOURCES

Considering evaluation upfront can help to ensure appropriate resourcing is in place.

Evaluations should be scaled to the program. A guide of 1–10% of the cost of the program is often described as appropriate, although this may vary for many reasons, including the focus and design of the evaluation.[7]

Evaluations can also be scaled to the resources available. Being clear about what resources are available for an evaluation will help to ensure the development of a feasible plan for the evaluation, focused on what is most important.

Evaluations can be conducted by an evaluation team within an organisation, or by an external evaluation team, so resources include:

- the staff resources (including knowledge, skills, and capacity) available for conducting the evaluation
- the money available for contracting an external evaluation team to undertake the evaluation

[7] NSW Treasury (2023). Policy and Guidelines: Evaluation. https://www.treasury.nsw.gov.au/documents/tpg22-22-policy-and-guidelines-evaluation

- the money available to remunerate advisory committees (where participants, such as those with lived and/or living experience, and community members are not employed by the organisation) and to provide payments to service users to recognise the time involved in their contribution to evaluation
- previous evaluations that can be drawn on to inform your evaluation design and compare your findings against.

Even if an external evaluation team is engaged, staff will still need time to:

- develop the request for proposal and review proposals
- meet with the evaluation team to brief them
- review the proposed evaluation plan
- work collaboratively with the evaluation team to support the evaluation process
- providing access to de-identified data and making connections to stakeholders, including service users
- review evaluation reports and implications
- consider how they will use findings.

4.2 DECIDING WHETHER AND WHEN TO EVALUATE A PROGRAM

As evaluations take time and resources, there is a need to weigh the cost of evaluation against the risk of not evaluating. Various government agencies have their own criteria for prioritising programs to be evaluated.

In the context of mental health and suicide prevention programs, criteria for prioritising programs and aspects of the system for evaluation include the following:

- **Funding/ investment**: programs with higher investment or ahead of additional investment, or scale-up or decision-making about ongoing funding
- **Risk**: programs that are high profile or sensitive, or where monitoring data, service user or workforce feedback suggest things are off-track or not working as intended
- **Available evidence**: programs with limited existing evidence or with limited existing evidence for the target group
- **Equity and priority populations**: programs that seek to address barriers to accessing services and/or target priority populations for mental health and suicide prevention
- **Service integration**: evaluations of programs aimed at improving integration across service boundaries or improving systems.

4.2.1 TIMING OF EVALUATION

If an evaluation is being conducted, it should be timed to align with decision-making requirements, for example, decisions about scale-up or further funding.

Some government agencies require evaluations be completed before funding renewal. Depending on the initial funding period, further evaluation might be required beyond this to enable a focus on medium- and longer-term outcomes.

The need for an evaluation might also be identified if monitoring data suggests the program is not working as intended, as evaluation might be required to identify why this is and what can be done.

4.3 ESTABLISHING GOVERNANCE ARRANGEMENTS

For evaluations involving multiple levels of government or government agencies, it is important to agree on clear roles and responsibilities. This may also include consideration of cost sharing of evaluations.

There is also a need to specify the role of funded service providers in evaluation, for example, providing access to de-identified data, inviting service users to participate in evaluation.

If an external evaluation provider is contracted, there is also a need to clearly outline the role of the evaluation team in an agreed scope of works.

4.3.1 OVERSIGHT OR ADVISORY COMMITTEES

Clear oversight arrangements should be established for evaluations, such as steering committees and/or advisory committees.

These committees can have an important role in ensuring evaluation designs are both robust and feasible, and in building stakeholder acceptance of evaluation recommendations.

An advisory or steering committee should have clear terms of reference. Their role in evaluation governance could encompass:

- informing the request for quote or tender for the evaluation (if being undertaken externally)
- reviewing proposal/s for the evaluation
- reviewing, contributing to, and potentially approving detailed evaluation plans
- providing lived and/or living experience and expertise through designated roles, and/or facilitating lived and/or living experience partnerships and engagement
- Aboriginal and Torres Strait Islander governance and cultural safety of the evaluation
- inclusion and safety for diverse communities, including culturally and linguistically diverse communities
- supporting the evaluation to connect with key stakeholder groups
- reviewing, contributing to, and potentially accepting evaluation reports
- informing key policy and program implications of evaluation findings
- supporting the dissemination of findings to key decision-makers.

The terms of reference should make clear how decisions about the evaluation are made and whether the committee has an advisory or decision-making/ sign-off role.

Where different levels of government or government agencies are involved, advisory committees should include representatives from all relevant agencies to ensure local contextual factors are considered in the evaluation design and interpretation of findings. This

can also provide the opportunity for additional questions, and data collection to be considered in different jurisdictions.

Consideration should be given as to whether there is one committee that includes funders, lived and/or living experience representatives, representatives of priority groups and/or external experts, or if there is more than one committee.

4.3.2 LIVED AND/OR LIVING EXPERIENCE LEADERSHIP

PARTICIPATION IN ADVISORY COMMITTEES

Where people with lived and/or living experience are included in oversight or advisory committees, it is important to consider:

- clarity of the role
- how power dynamics will be managed, to ensure people with lived and/or living experience of mental ill-health and/or suicide can safely and meaningfully contribute
- involving several lived and/or living experience representatives
- intersectionality and including people with certain lived and/or living experience expertise relevant to specific topics/ areas as appropriate
- any additional supports that might be needed to facilitate meaningful engagement, such as briefing and/or debriefing options and peer support
- appropriate remuneration, particularly as other governance committee members would usually be doing this as part of a paid role.

PARTICIPATION AND LEADERSHIP IN EVALUATIONS

Given the importance of lived and/or living experience leadership, evaluation teams should consider engaging team members with lived and/or living experience of mental ill-health and/or suicide, and those with experiences a family member, carer and/or support person (referred to throughout as people or team members with lived and/or living experience).

Involving lived and/or living experience leadership in evaluation teams can:

- demonstrate best practice and further spread lived and/or living experience leadership
- enhance the appropriateness of evaluation methods
- inform and improve recruitment strategies to engage people accessing the service in the evaluation
- help people accessing the service to feel comfortable providing feedback to the evaluation
- enable a more person-centred and lived and/or living experience-led interpretation of evaluation findings
- improve communication of results and implementation of recommendations.[8]

[8] Krysinska, K., Ozols, I., Ross, A. et al (2023). Active involvement of people with lived experience of suicide in suicide research: a Delphi consensus study. *BMC Psychiatry* 23, 496. https://doi.org/10.1186/s12888-023-04973-9

Appendix 3 provides more detailed guidance around engaging team members with lived and/or living experience so that organisations are conscious of the unique experience and insights of people with lived and/or living experience and a best practice approach. This includes appropriate renumeration and training to support their participation in governance or committee structures.

Resources

Hodges, E., Leditschke, A., Solonsch, L. (2023). *The Lived Experience Governance Framework: Centring People, Identity and Human Rights for the Benefit of All*. Prepared by LELAN (SA Lived Experience Leadership & Advocacy Network) for the National Mental Health Consumer and Carer Forum and the National PHN Mental Health Lived Experience Engagement Network. Canberra: Mental Health Australia.

Lived Experience Digital Library: The Lived Experience Leadership Digital Library is an initiative of the National Mental Health Consumer and Carer Forum and the National Primary Health Network Mental Health Lived Experience Engagement Network with support from Mental Health Australia.

4.3.3 ABORIGINAL AND TORRES STRAIT ISLANDER GOVERNANCE

In evaluations of programs for Aboriginal and Torres Strait Islander communities, or that involve analysis of outcomes for Aboriginal and Torres Strait Islander communities, it is important that appropriate Aboriginal and Torres Strait Islander Governance is established from the outset to ensure the evaluation is culturally safe.

To effectively evaluate with Indigenous communities, it is important to think about what you are doing but also why and how. This includes understanding the ongoing impact of colonisation. It is crucial to ensure engagement of representatives from the communities involved in the evaluation process. Evaluation commissioners need to build time into the evaluation for establishing and maintaining this engagement. Evaluators need to invest time upfront to build relationships with communities and establish governance arrangements.

The role of an Aboriginal and Torres Strait Islander Governance committee must be negotiated with the committee, but will likely include:

- informing stakeholder and community recruitment and engagement processes
- supporting relationship development with stakeholders and communities
- reviewing data collection approaches to ensure they are culturally respectful, and take into account ways of doing business
- supporting effective Indigenous Data Governance to ensure Indigenous Data Sovereignty
- reviewing analysis and interpretation of data to ensure interpretation is in context.

Appropriate remuneration should be given to contributing committee members who are not in paid roles, to reduce power imbalance, increase participation, and ensure equity.

Resources

Productivity Commission (2020). Indigenous Evaluation Strategy.

5. DECIDING THE FOCUS OF EVALUATION

This chapter describes key focus areas for evaluation – mapping different ways of describing these to enable synthesis across evaluations of different programs. It also provides a menu of key evaluation questions mapped to these focus areas, to support consistency across evaluations and a systems-level view across evaluations.

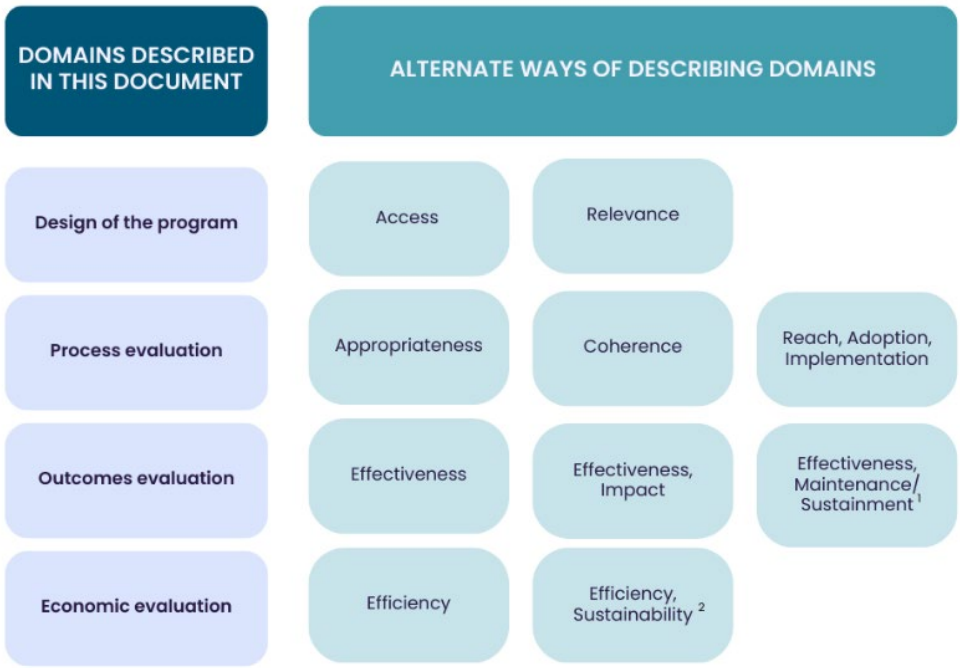
5.1 FOCUS AREAS FOR EVALUATION

Organisations and evaluators have many ways of describing the focus of evaluations. To support more consistency between evaluations of mental health and suicide prevention programs, and facilitate comparison of findings across programs, we have mapped the areas of focus commonly used (see Figure 3 below).

For this Framework, we have used the terms program (or initiative) design, process evaluation, outcomes evaluation, and economic evaluation to support consistency with government evaluation strategies.

Depending on when an evaluation is rolled out, it may be more or less relevant to focus on each of these domains. For example, a focus on design in the early stages can help inform refinements to the design, while a more comprehensive economic evaluation will likely be more feasible at a later stage of implementation.

FIGURE 3. FOCUS FOR EVALUATION



Source: 1 RE-AIM Framework (2023). https://re-aim.org/ ; 2 OECD DAC Network on Development Evaluation (n.d.) Evaluation criteria. https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

5.1.1 DESIGN OF THE PROGRAM

Evaluations that focus on the design of the program will focus on how a program was developed and/or whether it's meeting the intended needs and supporting the intended outcomes.

Programs should be informed by the specific needs and experiences of people who will be using them (people with lived and/or living experience of mental ill-health and/or suicide, along with families, carers and support people), and people who will be providing the service (the workforce). An appropriate understanding of needs – as well as the strengths in communities that can be built on – can help ensure programs focus on the right things.

If an evaluation team is involved **before the rollout of a program**, they can undertake a needs assessment: a systematic process used to identify, analyse, and prioritise the specific requirements, gaps, and challenges within a particular context or community and ensure the program is tailored to meet those needs or responsive to them.[9]

When a program is **in early stages of implementation**, an evaluation focused on the design might consider:

- how the needs to be addressed were identified
- the appropriateness of the design process
- the fit between the design and existing research evidence
- how the program was tailored to and reflects the needs of different communities where it is being implemented
- stakeholder perspectives on the appropriateness of the design.

If the program was co-designed or co-produced, it may be appropriate for the evaluation to include the co-design or co-production process.

5.1.2 PROCESS EVALUATION – IMPLEMENTATION

It is critical to understand the extent to which a program is **implemented as intended** before assessing outcomes. Understanding what is working well and what is not can help inform continuous improvement and increase the likelihood of success.

Understanding what has been critical to enable implementation and the barriers to implementation can inform consideration of further rollout or scale-up to support success.

A process evaluation should consider:

- whether a program is being implemented as intended
- whether and why adjustments were made
- consistency of implementation across locations (where there are multiple delivery locations)

[9] Smart, J. (2019). Needs assessment. Southbank: Australian Institute of Family Studies. https://aifs.gov.au/resources/practice-guides/needs-assessment#part-one:-defining-needs-and-needs-assessment

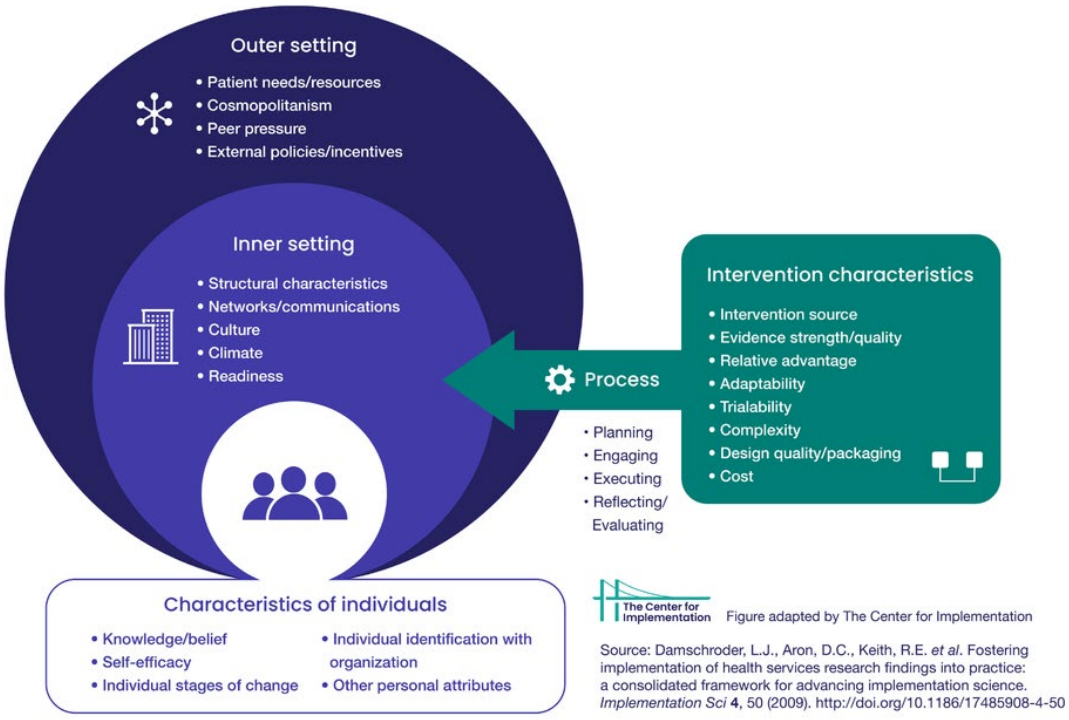
- whether the program is still the most appropriate response
- enablers and barriers to implementation.

A process evaluation may also consider the **appropriateness** of the program for the people for which it was designed, especially if this was not considered as part of an evaluation focused on the design of the program.

What "implemented as intended" means will differ between programs – from those using existing evidence-based models, where the model is expected to be implemented with "fidelity" (that is, consistent with the original model), to those that are co-produced, where iterative development and local variation are expected and need to be detailed and understood.

The *Consolidated Framework for Implementation Research* (Figure 4) can help to classify enablers and barriers to implementation. It is an organising Framework, based on the learning from over 500 published sources across many disciplines.[10]

FIGURE 4. CONSOLIDATED FRAMEWORK FOR IMPLEMENTATION RESEARCH



Source: CFIR Research Team-Center for Clinical Management Research (2022). Consolidated Framework for Implementation Research.

[10] Damschroder, L., Aron, D., Keith, R., Kirsh, S., Alexander, J., Lowery, J. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Science 7(4):50. https://doi.org/10.1186/1748-5908-4-50

5.1.3 OUTCOMES EVALUATION – EFFECTIVENESS

The terms **outcomes** and **impact** are used differently. In some cases, 'outcomes' is used to describe the more direct changes that result from an activity, and 'impacts' to describe the longer-term changes; in others, the reverse.

Outcomes evaluation examines if and how a program is leading to intended changes – which might be for individuals accessing the service, the workforce, and/or the broader system. It may include the extent and/or longevity of change, and the degree to which the program has contributed to the observed changes.

Outcomes evaluation should also identify if there have been any unintended outcomes – and the equity of outcomes. Outcomes evaluations can also assess the extent to which the program remains appropriate.

Understanding outcomes is important to understanding the value of a program. Depending on the timeframe of the evaluation, outcomes may be considered over the short, medium, or long term.

Evaluations of mental health and suicide prevention programs should consider – as relevant to the specific program – how they contribute to:

- improving equitable access to services and systems for the Australian population or the particular target group
- improving the mental health and wellbeing of the Australian population or the particular target group
- reducing suicide, suicidal distress, distress, and self-harm
- improving coordination of services and systems for all communities and groups
- improving physical health and life expectancy for people living with mental health conditions, and for those experiencing suicidal distress
- improving quality, safety, and capacity in the Australian mental health and suicide prevention system
- growing and supporting the mental health and/or suicide prevention workforce.

5.1.4 ECONOMIC EVALUATION – VALUE FOR MONEY

The purpose of conducting an **economic evaluation** – or assessing value for money in evaluation – is to inform decision-makers, stakeholders, and funders about whether a particular program represents a wise and **cost-effective** use of resources. This can help to inform choices about resource allocation. However, resource allocation can be informed by all evaluation types, not just economic evaluations. For example, investment decisions can also be informed by needs assessments, including mapping existing services to avoid duplication.

By assessing efficiency, evaluators can identify areas where resources are being used sub-optimally and recommend changes to enhance the program's efficiency and impact. However, efficiency cannot be considered in isolation from equity and other considerations, such as service quality and long-lasting outcomes.

5.2 KEY EVALUATION QUESTIONS

Effectively defining key evaluation questions (KEQs) – aligned to the focus areas for evaluation – ensures the evaluation provides the information required for intended users and decision-making.

Table 1 below provides a menu of system-level key evaluation questions mapped to program-level evaluation questions.

Not all evaluations of mental health and suicide prevention programs would be expected to answer all program-level questions. Some questions are inappropriate to ask at certain stages of a program, and some will not be relevant to a particular program. Also, it is not possible to answer all questions in a single evaluation. A good guide is to use a small number of KEQs – about three to seven questions, with some more detailed sub-questions under these. Individual evaluations may also have additional specific questions to be answered to inform local decision-making that do not appear in this menu.

The menu of program-level evaluation questions is intended to support some consistency across evaluations to facilitate the synthesis of evaluations at a systems level. The menu of program-level questions includes consideration about how programs interact with and influence mental health and suicide prevention systems, noting that individual programs can be influenced by system level factors or aim to influence system level factors.

However, it is important to note that any systems-level view produced based on a synthesis of program evaluations will have limitations, because not all programs are evaluated and not all evaluations are available. There are also some system-level questions that will be best answered through specific system-level studies – for example, questions about who is not being reached and why and a more complete picture of how services interact.

TABLE 1. KEY EVALUATION QUESTIONS AT THE SYSTEM AND PROGRAM LEVEL

Domain	System level	Program level
Program design	• To what extent do programs address identified needs? Where are there still gaps in the system? • How appropriate have the design processes been? How have people with lived and/or living experience and other stakeholders been involved in the design? • What opportunities remain for more nationally consistent mental health and suicide prevention programs?	• To what extent is there a need for this program in this context/ within the system? • How does the program align with the existing evidence base? • How appropriate was the design process? • How were people with lived and/or living experience and other stakeholders involved in the design?

Process evaluation – implementation	• To what extent is there a whole-of-government approach to coordinated mental health and suicide prevention, early intervention, treatment, aftercare, and postvention supports?
	• To what extent are mental health and suicide prevention programs reaching intended target groups, including priority populations?
	  ○ Who is being reached and why?
	  ○ Who is not being reached and why?
	• To what extent are mental health and suicide prevention programs enhancing access to supports and services, including for priority populations?
	  ○ What enables access?
	  ○ What are the barriers to access?
	  ○ To what extent are there appropriate connections between programs to support continuity of care?
	• What lessons can be learnt for implementing similar programs in future?
		• To what extent has the program been implemented as intended?
		  ○ What were the enablers and barriers to implementation?
		• To what extent has the program reached the intended target groups?
		  ○ Who is not being reached and why?
		  ○ Is this response appropriate?
		• What is the experience of people accessing the service?
		  ○ To what extent is the program acceptable to people intended to access it?
		• What is the experience of staff who deliver the program?
		  ○ Is staff selection appropriate?
		  ○ To what extent do staff receive sufficient training and support?
		• To what extent is the program considered safe and of high quality?
		• How does the program fit within the broader system?
		  ○ Are there appropriate connections and referral pathways to support continuity of care and holistic care?

Domain	System level	Program level
		• Would it be appropriate to scale-up or apply this program in other contexts and what would be required to support this?

Outcomes evaluation – effectiveness	• To what extent are mental health and suicide prevention programs improving the mental health and wellbeing of the Australian population?
	  ○ To what extent are the programs improving the mental health and wellbeing of priority populations?
	• To what extent are the programs contributing to a reduction in the prevalence of suicide, suicidal distress, and self-harm?
	  ○ To what extent are the programs reducing the prevalence of suicide, suicidal distress, distress, and self-harm among priority populations?
	• To what extent are the programs contributing to improvements in physical health and life expectancy for people living with mental health conditions and for those experiencing suicidal distress?
	  ○ What have been the outcomes for priority populations?
	• What have been the important factors in achieving these outcomes?
	• What improvements have there been at the systems level?
		• To what extent has the program achieved its intended outcomes?
		  ○ What is the value of these outcomes to stakeholders?
		  ○ What have been the important factors in achieving these outcomes?
		  ○ Were there any differences among people who access programs and services?
		  ○ Is this response still appropriate?
		• To what extent is the program contributing to system improvements? (where this is an aim)

Economic evaluation – value for money	• To what extent is funding providing value for money to governments and the Australian community? • To what extent is funding supporting a balanced and integrated mental health and suicide prevention system for all communities and groups? •	• How well have the resources been used? • To what extent is the program cost-effective/ does the program provide value for money? • To what extent has funding amount, longevity, and stability impacted on service capacity and quality? • To what extent does the program have benefits that are sustained?

6. DESIGNING A FIT-FOR-PURPOSE EVALUATION

Once the focus of the evaluation and key evaluation questions have been defined, you can choose an appropriate evaluation approach, design, and methods.

6.1 CONSIDERATIONS TO INFORM EVALUATION APPROACHES AND DESIGNS

There is no one right approach, design, or method for evaluation of mental health and suicide prevention programs. An appropriate evaluation approach and design is one which reflects:[11]

- key evaluation questions to be answered
- the scale of the program
- the resources
- the time available
- the nature and type of program, and existing evidence about the type of program (which may reduce the need for some new data collection)
- the stage of development of the program (when a program is being established it is more appropriate to focus on design and implementation, while you can more effectively capture outcomes of a mature program)
- availability and completeness of existing data (good quality data, collected as part of the program, may reduce the additional data that needs to be collected as part of the evaluation)
- contextual factors (what evaluation design is most appropriate for this program, with this group of people, in this context)
- ethical considerations and appropriateness for priority populations
- whether the program appears to be on track or not.

This chapter focuses on approaches to evaluation and impact evaluation designs you might consider, depending on these factors. But these factors will also influence the scale of your evaluation, the methods (considered in the next chapter), and whether the evaluation is conducted internally, externally or through a hybrid delivery approach. See Figure 5 for questions which may inform choices about design or approach.

[11] Patton, M.Q. (2014). Week 47: Rumination #3: Fools' gold: the widely touted methodological "gold standard" is neither golden nor a standard. https://www.betterevaluation.org/en/blog/fools_gold_widely_touted_methodological_gold_standard

FIGURE 5. QUESTIONS TO INFORM EVALUATION APPROACHES AND DESIGNS

Is program design established or expected to evolve?	• If the design will evolve, consider developmental evaluation to support this process. • If the design is established but hasn't been tried with this particular target group/s, you might consider a realist evaluation – understanding the mechanisms of change. • If the design is more established, you might also consider rubrics in which stakeholders agree on set criteria and standards against which to judge an initiative.
Is this program expected to engage with different target groups?	• If the design is different for different locations and target groups, you might consider case studies to deeply understand local differences, the rationale for these and how they are working. • If the program is expected to work differently for different target groups, you might consider a realist approach to understand what works for whom in what circumstances and why.
What stage is your program at?	• If it is early in the development of your program, it may be too early for robust outcomes and economic evaluation. • However, it is worth considering outcomes measurement early on if your intention is to understand impact through a randomised-controlled design or quasi-experimental design.
Is your program aiming to support changes in the system?	• If it is consider, a systems evaluation approach. • Even if it isn't, it might be worth considering how the outcomes for people in the program might be affected by factors in the system, such as availability of complementary services and referral pathways.
What is the completeness and quality of your existing data?	• If you routinely collect feedback and outcomes data from service users, this would reduce the scale of additional data collection required.
What's the scale of your program?	• For smaller scale programs, you may rely more on existing data collected as part of program delivery. • For larger scale programs, additional data collection may help you better understand the most valuable parts of the investment. • In larger scale evaluations, with larger client numbers, use of linked data may be more feasible. • If your program is actually multiple initiatives, you might consider rubrics, which would enable you to compare across different initiatives using agreed criteria.
How much budget do you have for evaluation?	• If you have a smaller budget, you might rely more on existing data or consider an empowerment evaluation approach. • A larger evaluation budget would allow for more in-depth data collection.
What is your timeframe?	• If you have a short timeframe, you will likely need to rely more on existing data because of the timeframes for ethics approval.

6.2 EVALUATION APPROACHES AND DESIGNS

6.2.1 CHOOSING AN APPROACH THAT IS FIT-FOR-PURPOSE

There are many different approaches to evaluation. Table 2 below outlines some key approaches to consider, their advantages and disadvantages, and when you might consider using them. Key considerations include the stage of development of the program, existing evidence about it, the scale of the program and resourcing for evaluation, and questions to be answered.

TABLE 2. EVALUATION APPROACHES

Approach	Description	Advantages	Disadvantages	Considerations for use
Developmental evaluation[12]	The evaluator facilitates regular data-based discussions with the program team about what is working and what isn't and what that means for practice. Evaluation can draw on a range of methods, as appropriate to the context. Methods are expected to evolve.	• Supports ownership and implementation of findings, centring accountability • Reflects the complex systems in which programs are implemented • Allows for understanding of how outcomes are supported to evolve • Can use different methods	• May make longer-term measurement more difficult if measures evolve • May be harder to synthesise with other evaluations because of targeted focus • Can be difficult to predict evaluation resources	• Suited for programs that are expected to continue to evolve through implementation • Suited to programs in complex systems, where outcomes or the processes for achieving them might be emergent (rather than predictable at the outset) • May not be appropriate where a level of independence is required for credibility of findings
Theory-based evaluation[13]	Use an explicit theory of change for a program to make conclusions about whether and how it contributed to observed outcomes. The terms theory of change and logic model are	• Governments commonly require logic models to be developed when a program is developed	• Logic models might not be all that logical – with huge leaps of faith between the identified	• A logic model and theory of change are often most useful when developed upfront to guide implementation, but they

[12] Patton, M.Q. (2011). Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. USA: Guilford Publications.

[13] Government of Canada (2022). Theory-Based Approaches to Evaluation: Concepts and Practices. https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/theory-based-approaches-evaluation-concepts-practices.html

Approach	Description	Advantages	Disadvantages	Considerations for use
	often used interchangeably and differently but should help to identify how change occurs (the mechanisms of change), assumptions, and risks.	• Can complement other approaches • Can help to identify gaps in existing evidence for the program that should be the focus of evaluation • Can help to identify where a program is on or off-track and highlight why this might be so • Can use different methods	activities and intended outcomes	can also be completed retrospectively • Can be difficult to represent emergence and non-linearity
Realist evaluation[14]	A realist approach recognises that the way a program 'works' will be different for individuals in different contexts. The focus of realist evaluation is not on answering 'what works?', but the more nuanced question of 'what works for whom in what circumstances and why?'	• By understanding who the program works for and how, a realist approach can help assess who to target and identify what matters for future implementation • Can help to identify whether a program in one setting might be able to be successful in another setting	• Can be difficult to operationalise, meaning some evaluations claiming to have a realist focus only identify where outcomes differ for different groups • Results might be more abstract	• Appropriate for programs in complex contexts • Appropriate for understanding how a program may work differently for different priority populations or communities

[14] Pawson, R. & Tilley, N. (2004). Realistic evaluation. British Cabinet Office. http://www.communitymatters.com.au/RE_chapter.pdf.

Approach	Description	Advantages	Disadvantages	Considerations for use
		• Can be used alongside a theory-based approach • Can use different methods		
Rubrics[15]	A rubric establishes an agreed set of criteria against which to assess the program, and performance standards for each criterion based on a shared understanding.	• Can be useful if you want to get people on the same page about what matters and how to assess what is of value • Rubrics allow for explicit judgements about the quality, value, or importance of the program based on an agreed set of criteria and performance standards. They make evaluative reasoning explicit • For more complex rubrics, criteria can be given different weights depending on their level of importance	• Different stakeholders may value different criteria differently	• Useful to systematically rate similar programs, for example, the same program delivered in many different locations or a suite of programs targeting shared outcomes • Allows for different stakeholder values to inform the evaluation

[15] Better Evaluation (2012) Creating Rubrics. https://www.betterevaluation.org/tools-resources/creating-rubrics

Approach	Description	Advantages	Disadvantages	Considerations for use
Systems evaluation	There is no one accepted approach to systems evaluation. Renger[16] has developed one approach that involves: • Defining the system, subsystems and the relationships between them. Includes: role of intermediaries and their relationships and interactions with other components of the systems; and the attributes of systems or subsystems, including skills, commitment of leadership, infrastructure, and organisational culture. • Assessing system efficiency. Evaluating feedback mechanisms and determining the extent to which subsystem processes and	• Systems thinking can help to understand the relevant structures, patterns of behaviours and events needed for change • Speaks to stakeholder interest in a systems perspective for mental health and suicide prevention programs and the importance of interrelationships between services and systems to outcomes	• Not a universally agreed approach • May be harder to synthesise with other evaluations because of targeted focus	• Appropriate where a system or systems are the key focus for analysis

[16] Renger, R. (2015). System evaluation theory (SET): A practical framework for evaluators to meet the challenges of system evaluation. Evaluation Journal of Australasia. 15. 16-28. http://dx.doi.org/10.1177/1035719X1501500403

Approach	Description	Advantages	Disadvantages	Considerations for use
	attributes are aligned to a common goal. • Monitoring and evaluating systems outcomes: Monitoring indicators specific to the problem conditions targeted by the system and subsystem, without seeking to generalise to other systems.			
Empowerment evaluation[17]	In an empowerment evaluation, program staff and community members are in control and the evaluator acts as a critical friend (or coach) – someone who believes in the program but is able to ask the critical questions to ensure an honest reflection on the evidence. Different practices have evolved, but the original includes three steps: 1. Establish the mission 2. Take stock of current status (against key	• Supports leadership by key stakeholders, so may be appropriate to support lived and/or living experience leadership • Supports action • Builds capability of program staff and others involved in the evaluation	• May make longer-term measurement more difficult if measures evolve • May be harder to synthesise with other evaluations because of targeted focus • Empowerment evaluation has been criticised within the evaluation community with some questions raised about whether it is	• May be particularly appropriate for smaller funded services/ programs that have limited experience with evaluation – who it was developed for • May not be appropriate where a level of independence is required for credibility of findings

[17] Fetterman, D., Kaftarian, S.J., & Wandersman, A. (2015). Empowerment evaluation: Knowledge and Tools for Self-assessment, Evaluation Capacity Building, and Accountability, Second Edition. USA: SAGE Publications Inc.

Approach	Description	Advantages	Disadvantages	Considerations for use
	activities to achieve the mission) 3. Plan for the future (identify strategies and methods for tracking success).		evaluation and how it deals with bias	

6.2.2 CHOOSING AN OUTCOMES EVALUATION DESIGN

When it comes to outcomes evaluation, there are a range of potential evaluation designs that may be more or less appropriate in different contexts (as illustrated in Table 3 below).

TABLE 3. EVALUATION DESIGNS

Design	Description	Advantages	Disadvantages	Considerations for use
Experimental designs	Experimental designs: include randomised controlled trials (RCT), cluster RCTs, stepped wedge designs. In experimental designs, participants are randomly assigned to a group that receives the health 'intervention' or a comparison group that does not receive the intervention. Data for each group are collected before and after the intervention. In theory, randomisation means the groups will as similar as possible at the outset of the study, though the likelihood of this increases with increasing sample size	• When ethical and feasible, well-designed experimental studies can establish a *cause-and-effect relationship*. • Experimental designs control for selection bias.	• Experimental designs are not always possible, ethical or practical in mental health and suicide prevention. • Experimental designs can be expensive to implement. • May be subject to challenges, such as loss to follow-up.	• May be difficult for more complex interventions that require flexibility in delivery. • There are a range of ways to facilitate randomisation[18], which expand the opportunity to consider experimental design • Randomise a group rather than an individual. For example, randomising regions. • Randomise in stages. For example, a staggered rollout. • More feasible for larger scale initiatives.
Quasi-experimental designs	Quasi-experimental designs involve the identification of a comparison group as similar as	• Can be used when it is not ethical to randomly assign people to a treatment group.	• It is likely less feasible to use this approach for programs delivered by community	• Can be used when it is not possible to randomly assign people to different groups.

[18] Department of the Treasury. (2023). Randomised controlled trials. https://evaluation.treasury.gov.au/toolkit/randomised-controlled-trials

Design	Description	Advantages	Disadvantages	Considerations for use
	possible to those accessing the program.		organisations, where it can be more complicated to link to administrative data sets that would enable the establishment of a comparison group.	• May be less feasible for initiatives delivered through NGOs.
Non-experimental designs	Non-experimental designs are those which do not have a comparison or control group.	• Often more feasible to implement, particularly for community-based programs that cannot draw on administrative data sets to create a comparison group • May help address context and contribution in complex programs	• They are often not considered as robust as experimental or quasi-experimental designs.	• Can be used when it is not possible to randomly assign people to different groups. • Can be used when linkage to administrative data to enable construction of a comparison group is not possible.

In an evaluation of a small-scale program, the use of a **non-experimental design**[19] will often be more feasible. There are a number of ways to increase the robustness or credibility of a non-experimental design, for example:

- Contribution analysis, which aims to make credible causal claims by assessing whether the program is based on a reasoned theory of change, the program was implemented as intended, the intended outcomes occurred, and other factors influencing the program were assessed and were either shown not to have made a significant contribution or, if they did, the relative contribution was recognised.[20]
- Qualitative comparative analysis: comparing the configurations of different cases to identify the components that produce specific outcomes.
- Process tracing: case-based approach to causal inference focused on the use of clues within a case (causal-process observations, CPOs) to adjudicate between alternative possible explanations.
- Comparative case studies[21]: involve analysis and synthesis of the similarities, differences and patterns across two or more cases with a common goal, in a way that produces knowledge that is easier to generalise about causal questions.

Useful references

Various government departments have their own guidance on and templates for evaluation designs. Other examples include:

Experimental designs
- *Better Evaluation: Randomised Controlled Trial*
- *Department of the Treasury: Randomised Controlled Trials*

Quasi-experimental design
- *Better Evaluation: Quasi-experimental designs*
- *NSW Treasury: Outcome Evaluation Designs*

Broader considerations
- *NSW Government Centre for Epidemiology and Evidence, Study Design for Evaluating Population Health and Health Service Interventions: A Guide*
- *Better Evaluation: Impact Evaluation*

[19] Rogers, P., Hawkins, A., McDonald, B., Macfarlan, A. & Milne, C. (2015). Choosing appropriate designs and methods for impact evaluation. Department of Industry innovation and Science. https://www.betterevaluation.org/sites/default/files/choosing_appropriate_designs_and_methods_for_impact_evaluation_2015.pdf

[20] Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation,* 18(3) 270–280.

[21] Goodrick, D. (2014). Comparative Case Studies. Italy: UNICEF Office of Research. https://www.unicef-irc.org/publications/754-comparative-case-studies-methodological-briefs-impact-evaluation-no-9.html#:~:text=Comparative%20case%20studies%20involve%20the,work%20or%20fail%20to%20work.

6.2.3 MIXED METHODS

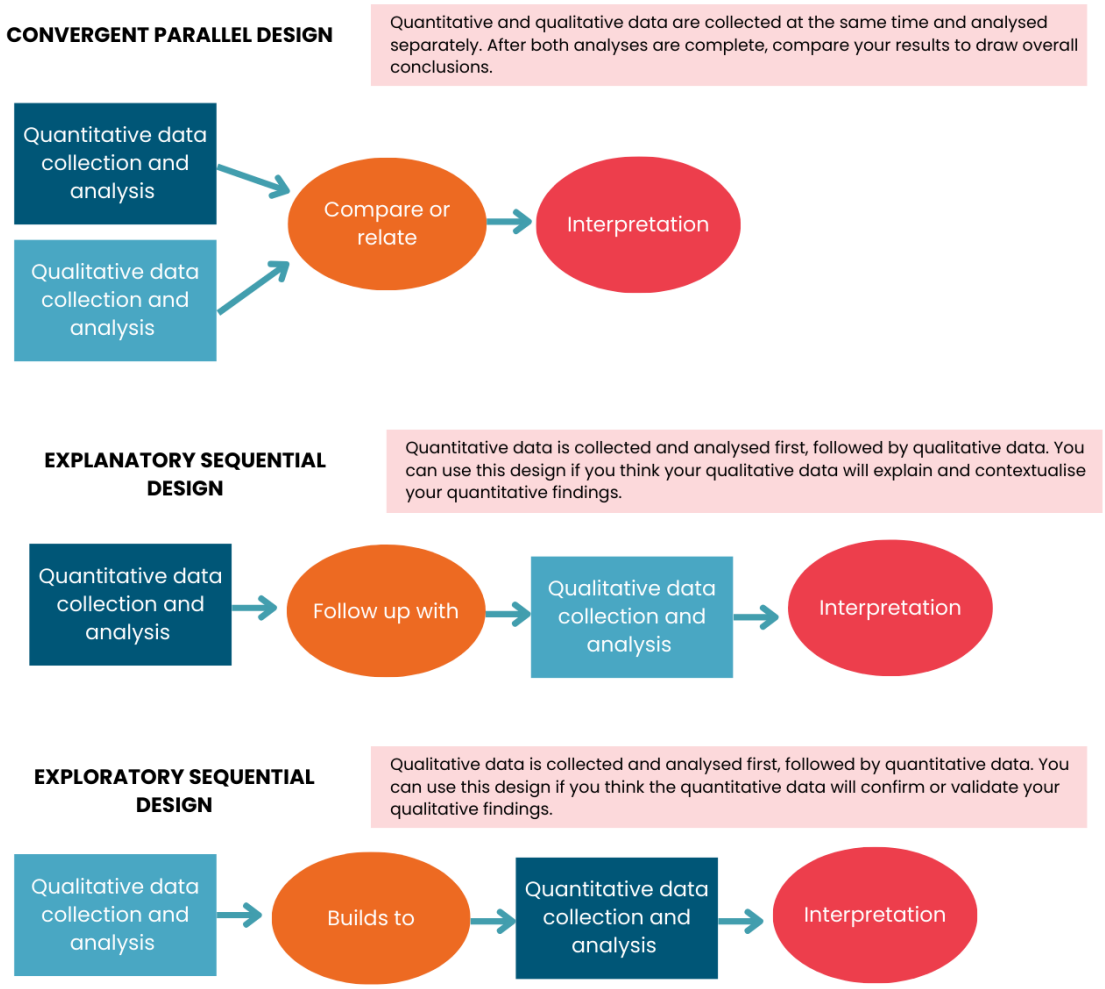
A mixed methods design – analysing a combination of quantitative and qualitative data to answer key evaluation questions – has benefits. Quantitative data can provide a statistically reliable and robust measure of behaviours, opinions, and preferences of a population, while qualitative data helps to understand patterns behind the quantitative data, complexity, and context.

Key features of a mixed methods approach are:

- **triangulation**: to corroborate findings between both quantitative and qualitative data
- **complementarity**: to enhance and clarify results from one method with the results of another
- **development**: to use results from one method to inform the development and construction of the other method.

There are a range of ways to mix methods in evaluation. Figure 6 describes different ways of sequencing methods and the potential benefits of these.

FIGURE 6. MIXED METHOD DESIGNS



More about mixed methods can be found here: *Conducting mixed-method evaluations | Better Evaluation*

6.2.4 DESIGNING AN EVALUATION THAT IS CULTURALLY SAFE

It is important that evaluation is culturally safe for Aboriginal and Torres Strait Islander communities.

The Productivity Commission's Indigenous Evaluation Strategy[22] for evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people, specifies that:

- evaluations be undertaken in the areas, and address the issues, that are most important to Aboriginal and Torres Strait Islander people
- Australian Government agencies routinely consider the impacts of mainstream policies and programs on Aboriginal and Torres Strait Islander people
- Aboriginal and Torres Strait Islander people, organisations, and communities have the opportunity to decide how they want to be involved in evaluations
- non-Indigenous evaluators have the necessary knowledge, skills, experience, and awareness of their own biases to work in partnership with, and to draw on the knowledges of, Aboriginal and Torres Strait Islander people
- evaluation processes strengthen and support the evaluation capability of Aboriginal and Torres Strait Islander people.

The *Solutions that work: What the evidence and our people tell us,* Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report[23] recommends a framework for evaluating suicide prevention programs in Aboriginal and Torres Strait Islander communities, structured around developing a logic model for the program. It also recommends Participatory Action Research as one appropriate approach for future suicide prevention research in Aboriginal and Torres Strait Islander communities. This approach may also be appropriate for suicide prevention and mental health programs more broadly.

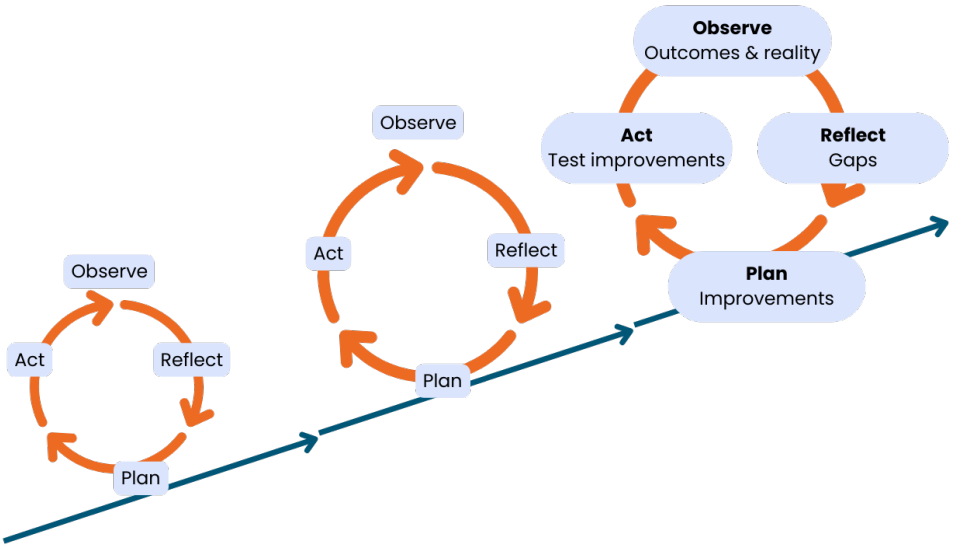
Participatory Action Research is an approach to research that aims to situate power in the research process with those who are most affected by a program. It engages those with a stake in the program – such as staff, people accessing the program, and community members – in defining the questions, data collection, and shaping actions. It is guided by

[22] Productivity Commission (2020). Indigenous Evaluation Strategy. https://www.pc.gov.au/inquiries/completed/indigenous-evaluation/strategy/indigenous-evaluation-strategy.pdf

[23] Dudgeon, P. et al (2016). Solutions that work: What the evidence and our people tell us: Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report. School of Indigenous Studies, University of Western Australia Centre for Child Health Research.

principles of social change, participation, empowerment, and collaboration.[24] [25] [26] [27] It commonly involves four steps (Figure 7 below), which are used in cycles to support change over time.

FIGURE 7. THE PARTICIPATORY ACTION RESEARCH CYCLE



Participatory Action Research shares some qualities with developmental evaluation and empowerment evaluation approaches, discussed in section 6.1 and 6.2. Table 4 below identifies advantages and disadvantages of the approach, and considerations for when it is appropriate.

[24] Rosier, K., Lohoar, S., Moore, S. & Robinson, E. (2015). Participatory Action Research. Australian Institute of Family Studies. https://aifs.gov.au/resources/practice-guides/participatory-action-research
[25] Pain, R., Whitman, G., Milledge, D., & Lune Rivers Trust. (2011). Participatory action research toolkit: An introduction to using PAR as an approach to learning, research and action. Durham: Durham University
[26] Boyle, M. (2012). Research in action: A guide to Participatory Action Research (Research Report). Canberra: Department of Social Services, www.dss.gov.au/sites/default/files/documents/06_2012/research_in_action.pdf
[27] Patton, M. Q. (2008). Utilization-focused evaluation (4th ed.). London: SAGE.

TABLE 4. CONSIDERATIONS FOR THE USE OF PARTICIPATORY ACTION RESEARCH

Advantages	Disadvantages	Considerations for when it is appropriate to use
• Centres Aboriginal and Torres Strait Islander leadership • Increases depth and understanding of issues, increasing relevance and reducing potential for misunderstanding • Increases community ownership and use of findings • Can help to overcome logistical barriers to data collection • Draws on stakeholder knowledge to help improve the program	• The stakeholders involved may not represent all stakeholder interests • Can be more time-consuming • Can be susceptible to researcher bias • Not as well recognised as quantitative approaches • Can produce a lot of data that may be difficult to manage	• Requires community leadership and interest in leading the process, not just using the findings • Useful for a program in development, where stakeholders may have already identified opportunities for improvement • May not be appropriate where more technical research methods are needed, or independence is required to give credibility to evaluation findings

Source: Rosier, K., Lohoar, S., Moore, S. & Robinson, E. (2015). Participatory Action Research. Australian Institute of Family Studies.

6.2.5 EVALUATING PEER WORK

Limitations of traditional evaluation approaches – evaluations focused on linear cause and effect relationships – have been identified for evaluating peer work for a range of reasons. This includes the focus on supporting personal recovery, which might not align with typical outcome measures, such as avoidance of health services or reduction in healthcare costs, as well as the challenges peer workers can experience working in biomedically oriented systems.

Towards a Meaningful Evaluation Framework for Peer Work from the Centre for Social Justice and Inclusion at the University of Technology Sydney proposes a focus on the following drawing on the Map of Meaning.[28]

- Reflect: Developing the inner self, including moral development, personal growth, or an authentic self
- Connect: Creating a sense of connection, belonging and solidarity across intersections of identity
- Respect: opportunities to make a difference, and serve others
- Express: Enabling people to express their potential, and to achieve goals and influence others.

This project also emphasises the importance of engaging people with lived and/or living experience in interpreting data.

[28] Lips-Wiersma, M. & Morris. L. (2018). The Map of Meaningful Work – A Practical guide to sustaining our humanity (Revised & updated 2nd ed.) London: Routledge.

7. CHOOSING APPROPRIATE METHODS AND MEASURES

7.1 EVALUATION METHODS

Once an appropriate evaluation design has been identified, the next step is to confirm the methods and measures to answer the key evaluation questions. The methods used in an evaluation will depend on the purpose of the evaluation, existing data, and the available time and resources. To reduce the burden on people accessing the service and on staff, evaluations should make the most of existing administrative data and use additional data collection to fill gaps in these.

Figure 8 identifies key methods for data collection, and considerations for when and how these are used.

FIGURE 8. DATA COLLECTION METHODS AND THEIR POTENTIAL USES

METHOD	HOW TO USE IT	WHEN TO USE IT
Desktop research/ document review	A systematic review of documents, which may include existing published documents, program guidelines and procedures. This may also include a review of existing data.	Design phase: To gather ideas and evidence when designing an initiative or developing a business case.
Literature review	A systematic review of similar programs run in other jurisdictions, reviews or evaluations of similar programs, relevant journal research articles.	Design phase: To gather ideas and evidence when designing an initiative or developing a business case.
Stakeholder interviews	Semi-structured interviews with a range of stakeholders (face-to-face, telephone, or video-conferencing). With permissions, interviews can be audio recorded to enable transcription and improve the accuracy of analysis.	Design, implementation, and outcomes phase: Stakeholders to be interviewed will be identified when refining the methodology of each evaluation.
Interviews or group interviews	Interviews or groups interviews with consumers, families and carers, and service providers will help answer access, appropriateness, and experience questions.	Implementation and outcomes phase: The details of this will be identified when refining the methodology of each evaluation.
Surveys	Surveys are useful to build background and baseline information. Survey design should be structured in order to ask the right questions of the right people to answer your key evaluation questions.	Implementation and outcomes phase: The details of this will be identified when refining the methodology of each evaluation.
Administrative data, linked data	As part of service delivery and to meet contractual requirements, service providers often collect and store client data, including socio-demographic characteristics, health or service needs, referrals made or received, types of services provided and frequency of access, and assessment results.	Monitoring, implementation, and outcomes phase: This data helps us to understand the people accessing and using a service, the nature of their interactions, and, in some cases, service costs and outcomes. Availability and strategies to ensure completeness of data need to be considered at the design stage.
Case studies or case stories	A case study focuses on the experience of a particular person, site, or project. Case studies and case stories can be particularly useful for understanding how different elements fit together and how different elements (implementation, context, and other factors) have produced the observed impacts.	Implementation and Outcomes phase. Data can be drawn from many sources. It is often in the use of case studies, vignettes and stories that policymakers really understand the experiences of consumers, carers and their families, and the implications for service design and delivery.

7.2 IDENTIFYING APPROPRIATE MEASURES

If evaluation has been considered upfront, and data about the program routinely collected, evaluators will be able to draw on this data for evaluation. The sections below outline typical measures that might be used for each of the focus areas for evaluation of mental health and suicide prevention programs.

7.2.1 DESIGN MEASURES

When the focus is on the design of the program, the measures to consider include the need for the program, how it aligns with existing evidence, and the design process.

THE NEEDS BEING ADDRESSED

To appropriately assess the design, there is a need to understand:

- the particular need to be addressed by the program
- the target group for the program and their needs (which may include evidence from population data about mental health and suicide prevention)
- alignment with other existing programs.

HOW THE PROGRAM ALIGNS WITH THE EVIDENCE

Assessment of the design should consider alignment with existing evidence, while noting that this may be more limited for innovative programs or initiatives addressing needs in new ways. Evidence should be considered broadly:

- how the program model aligns with research literature or previous evaluations of similar programs that identify particular program models or core components of programs known to be effective
- how the program incorporates practitioner knowledge and the insights of those with lived and/or living experience
- how the program reflects needs and preferences of people who may access the program.

In this process, it is important to note that there is likely to be more limited evidence about programs that work for some priority populations.

Factors to consider in the design of Aboriginal and Torres Strait Islander suicide prevention programs

Solutions that work: What the evidence and our people tell us, Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report[29] identifies success factors for Indigenous suicide prevention programs from a meta-evaluation of evaluated community-led Indigenous suicide prevention programs. These include factors associated with:

[29] Dudgeon, P. et al (2016). Solutions that work: What the evidence and our people tell us: Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report. School of Indigenous Studies, University of Western Australia Centre for Child Health Research.

- prevention (addressing community challenges, poverty, social determinants of health; cultural elements – building identity, social emotional wellbeing, healing, alcohol/ drug use reduction)
- primary prevention (such as awareness raising and tailored community gatekeeper training)
- clinical elements (access to counsellors/ mental health support, 24/7 availability, awareness of critical risk periods and responsiveness at those times, crisis response teams after a suicide/ postvention
- community leadership
- provider (partnerships with Aboriginal Community Controlled Organisations, employment of community members and peer workforce).

These could be a useful reference against which to assess the design of Aboriginal and Torres Strait Islander suicide prevention programs to identify if these elements are in place.

CO-DESIGN PROCESS

If a program was developed using a **co-design or co-production process**, then the evaluation of this program should also consider:

- whether the intended stakeholders, including people with lived and/or living experience, were engaged in the co-design or co-production process, and the diversity of representation
- whether the process was in line with best practice principles for co-design and co-production
- whether the resulting design reflected the inputs of the co-design or co-production process, and where input was not possible within the scope of the program this was clearly communicated.

Those using co-design approaches should consider building in feedback on the process as part of co-design, because this information can be difficult to collect retrospectively.

7.2.2 PROCESS EVALUATION MEASURES

A range of measures are needed to understand how a program is being implemented. Process evaluation measures can provide valuable insights about whether the program might need to be adapted to meet the needs of some population groups .

Many measures of implementation – such as who is being reached (or who is missing out) and what is being delivered – will be collected as part of program delivery and in required minimum data sets. Evaluators should look first to making the most of this existing data.

The **National Mental Health Performance Framework** (2020) also identifies some specific aspects of implementation that may be useful for mental health and suicide preventions to reference in a process evaluation.

These include the extent to which a program:

- is **accessible** – people can access care at the right place and the right time, within the context of different population needs and the affordability of care
- **reaches the people it is intended to reach** – the number, proportion, and representativeness of individuals who are able or willing to participate in a given program (and reasons why or why not)
- is **safe and of sufficient quality** – safety of care delivered to people accessing the service, as well as safety of carers and workforce
- is **appropriate or acceptable** – such as whether a program is person-centred, culturally appropriate, rights-based, trauma-informed, and recovery oriented; mental health consumers and carers are treated with dignity and confidentiality and encouraged to participate in choices related to their care
- is **implemented as intended** – fidelity, where there is an established model.

For a deeper understanding of implementation, for example, the reasons for differences between two or more program locations, evaluators will likely need to collect qualitative data to provide contextual information to help explain findings.

WHO RECEIVES (AND WHO MISSES OUT)

Indicators of accessibility are designed to ensure that people are obtaining care at the right place and the right time, taking account of different population needs and the affordability of care. Some measures to consider including in an evaluation to determine accessibility are:

- demographic and clinical characteristics of people who access programs and services, collected at episode level by service providers
- wait lists to determine need and accessibility
- acceptance and disengagement rates for the program
- availability of other relevant supports and services
- referral, triaging, and intake/ assessment procedures
- understand who is missing out and why.

Collection of demographic data is important to understanding the reach into priority groups – that is, who is accessing services and who is missing out. The appropriate collection of demographic data can also help services to provide culturally appropriate support, including for Aboriginal and Torres Strait Islander communities, and support tailored to priority populations.

WHAT SERVICES

Understanding what services are being received/ provided is important in determining if the program is being delivered as intended. Indicators about the services received may cover:

- data collected by service providers for each individual service event (e.g., date and type of service, duration)
- how well services are integrated around consumers (person-centred, recovery-oriented, trauma-informed models of service delivery).

FROM WHOM

Understanding who is delivering the services is important to understand the experience of people accessing the services.

These include:

- service provider and organisation characteristics
- details of organisation and mental health workforce delivering services, reported by the service provider.

Additional measures might include the nature of the workforce, including clinical and peer workforces, workforce diversity, the training provided to the workforce, and workforce experience. This information may also support the growing investment in peer workforces.

WITH WHAT EXPERIENCE

Experience indicators may include:

- experience directly collected from people accessing the service (with some standard data collection tools used for this)
- appropriateness of the service (provision of a person-centred, culturally appropriate, rights-based, trauma-informed, and recovery-oriented service, participation in choices related to care)
- safety of the service (such as patient reported incidents).

It can be difficult to understand why people drop out of a program, because it can be hard to reach them. But getting some information about this is important as it may relate to negative experiences or unintended consequences. If it is not possible to speak to individuals, look for patterns in the data about who is not engaging or who is dropping out and when, or look to existing research about barriers to access and engagement.

SERVICE CONNECTIONS

It is also important to understand integrated care and connections between services. Service connection indicators might include:

- data on referrals into and out of a service
- service user experience survey or interviews
- partner/ collaboration survey or interviews.

OTHER MEASURES

A useful focus for process evaluation, which will likely require qualitative data collection, is the identification of enablers and barriers to implementation, and how these have been leveraged and overcome. This is important to informing further rollout of a program, so that the conditions for success are in place.

7.2.3 OUTCOMES EVALUATION MEASURES

Identifying appropriate outcomes measures for a program is an important first step for an outcomes evaluation. This is important to improving the broader evidence base for mental health and suicide prevention.

Options for outcomes measures to use in evaluation include:

- client outcomes data collected by providers, often using standard instruments as part of minimum data sets
- administrative data, including linked data
- qualitative data about the outcomes valued by people accessing the service.

There are challenges evaluating the outcomes of mental health and suicide prevention programs. The rate or number of suicide deaths is unlikely to be an appropriate outcome measure for evaluations of individual suicide prevention programs, particularly in the short-term because of the small number of suicide deaths, the confounding factors that cannot be controlled for, and the lag that often exists between a service and outcomes due to the multifactorial nature of suicide and the diverse contributors to it. Measuring reduced suicide-related distress, in addition to increased resilience and help-seeking, is usually more appropriate.

STANDARD OUTCOMES MEASURES

Some standard outcomes measures are currently being collected through two mental health and suicide prevention minimum data sets:

- The *Mental Health National Outcomes and Casemix Collection* (NOCC) comprises a range of clinician and consumer rated measures, being progressively implemented by states and territories. The NOCC measures contribute to the development of clinical practice, aiming to improve the quality of care for consumers of Australia's public sector mental health services.
- The *Primary Mental Health Care Minimum Data Set* provides the basis for PHNs and the Department of Health to monitor and report on the quantity and quality of service delivery, and to inform future improvements in the planning and funding of primary mental health care services funded by the Australian Government.

See Table 5 below for an overview of the measures, and Appendix 5 for more detail.

The use of this existing data for evaluations of programs required to report on minimum data sets will help reduce the burden of data collection for staff and people accessing services. It will also help to synthesise findings across evaluations to provide a systems-level view.

Programs not required to report on minimum data sets might also consider using these measures, where appropriate, to support the ability to synthesise across evaluations.

However, it is important to note that people accessing services have identified the limitations of standard measures for capturing the full value of a program, and suggest the importance of additional data collection – for example, qualitative questions to better capture the impact

of symptoms rather than the existence of symptoms and people's ability to manage their mental health.

When standard measures are used, they are more useful to both staff and people accessing supports when used to guide service delivery and conversations with the person accessing the service in an ongoing way, so they can track their outcomes.

TABLE 5. STANDARD OUTCOMES MEASURES

Source		Measure	Age groups	
NOCC	PMHC-MDS		Children & Adolescents	Adults
X		Health of the Nation Outcome Scale (HoNOS/ HoNOS65+)		X
X		Health of the Nation Outcome Scale for Children and Adolescents (HoNOSCA)	X	
X		Abbreviated Life Skills Profile (LSP-16)		X
X	X	Kessler Psychological Distress Scale (K10)		X
	X	Kessler-5 (K5; for Aboriginal and Torres Strait Islander people if considered more appropriate)		X
X		Behavior and Symptom Identification Scale (BASIS-32)		X
X		Mental Health Inventory (MHI-38)		X
X	X	Strengths and Difficulties Questionnaire (SDQ)	X	
X		Living in the Community Questionnaire (LCQ)		X
X		Factors Influencing Health Status (FIHS; not an outcomes measure but important to interpret outcomes data)		X

When set measures are not already collected as part of standard data collection, the figure below provides some guidance on the focus of the outcomes measures and whether they are mental health or suicide specific, to help funders, program staff, and evaluators identify their appropriateness for their program. Figure 9 also identifies some potential additional measures that are not captured in standard measures – focused on recovery and suicidality – that organisations may consider using for evaluation, where recovery or reducing suicidality are among their intended outcomes.

Appendix 5 provides further detail about all of the standard measures, noting that consideration also needs to be given to their appropriateness for different priority populations.

FIGURE 9. FOCUS OF EXISTING OUTCOMES MEASURES

	OUTCOMES MEASURES	OUTCOMES
MENTAL HEALTH SPECIFIC	• Depression, Anxiety and Stress Scale (DASS) • Strengths and Difficulties Questionnaire (SDQ)* • Dialectical Behaviour Therapy Ways of Coping Checklist (DBT-WCCL) • Brief Family Distress Scale (BFDS) • Connor-Davidson Resilience Scale (CR-RISC) • Feel better • Kessler Psychological Distress Scale (K10)* • Kessler-5 (K5)* • Behaviour and Symptom Identification Scale (BASIS-32)*	• Behaviours • Resilience • Increase / decrease wellbeing • Distress
BROADER MEASURES	• Recovery Assessment Scale -Domains and Stages (RAS-DS) • Questionnaire about the Process of Recovery (QPR) • HaPI • Life Skills Profile (LSP-16)*	• Recovery appraisal • Seeking Help • Lifestyle • Functioning / not functioning
SUICIDE PREVENTION SPECIFIC	• Suicidal Ideation Attributes Scale (SIDAS) • Beck Scale for Suicidal Ideation (BSS)	• Reduction of attempting • Ideation reduction • Reduction of Emergency Department presentations • Exposure awareness^
QOL DOH SDOMH	• Physical Health • Living in the Community Questionnaire (LCQ)* • Work and Social Adjustment Scale (WSAS) • HaPI • Factors Influencing Health Staus (FIHS)* • Health of the Nation Outcome Scale (HoNOS/ HoNOS65+)* • Health of the Nation Outcome Scale for Children and Adolescents (HoNOSCA)*	• Overall functioning

^ NMH&WB collects exposure to suicide data as those exposed to suicide are higher risk of suicide. SEES has an evidence base for examining impact of this exposure.

*Measures included in either the NOCC or the PMHC-MDS.

APPROPRIATE OUTCOMES MEASURES FOR ABORIGINAL AND TORRES STRAIT ISLANDER COMMUNITIES

It is important that evaluations – including of mainstream services – engage Aboriginal and Torres Strait Islander service users to understand whether and how a service is working for them.

In many cases, standardised outcomes measures will not have been tested for their appropriateness for Aboriginal and Torres Strait Islander communities. Of the measures in the minimum data sets, the K-5 has more commonly been used with Aboriginal and Torres Strait Islander communities.

A more holistic consideration of outcomes is likely to better capture the value of a program. While the National Aboriginal and Torres Strait Islander Suicide Prevention Strategy is still under development, (as at October 2023), the National Strategic Framework for Aboriginal and Torres Strait Islander Peoples' Mental Health and Social and Emotional Wellbeing 2017-2023[30] provides some guidance on the domains of social and emotional wellbeing. These are connections to domains of social and emotional wellbeing outlined in Figure 10.

FIGURE 10. DOMAINS OF SOCIAL AND EMOTIONAL WELLBEING

Source:
Gee, G., Dudgeon, P., Schultz, C., Hart, A. & Kelly, K. (2013). Aboriginal and Torres Strait Islander Social and Emotional Wellbeing. In Working Together: Aboriginal and Torres Strait Islander Mental Health and Wellbeing Principles and Practice. Telethon Institute for Child Health Research/Kulunga Research Network, in collaboration with the University of Western Australia.

[30] https://www.niaa.gov.au/sites/default/files/publications/mhsewb-framework_0.pdf

The importance of collecting Aboriginal and Torres Strait Islander values-based social and emotional wellbeing and mental health outcomes measures, in combination with clinical measures, for Aboriginal and Torres Strait Islander mental health and suicide prevention services and programs is identified in the *Gayaa Dhuwi (Proud Spirit) Declaration: A companion declaration to the Whareratā Declaration*, as well as *Solutions that work: What the evidence and our people tell us*, Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report.[31] The Gayaa Dhuwi Proud Spirit Declaration Implementation Plan – not yet published – is expected to identify indicators with which evaluations could align.

The National Strategic Framework for Aboriginal and Torres Strait Islander Peoples' Mental Health and Social and Emotional Wellbeing 2017-2023 guidance notes that measures should be developed under the leadership of, and in partnership with, Aboriginal and Torres Strait Islander leadership bodies. Indicators should draw on Aboriginal and Torres Strait Islander understandings of how to measure mental health and social and emotional wellbeing, as well as general population measures including those from the Mental Health Management Performance Framework 2020. Indicators could include:

- social and emotional wellbeing
- mental health difficulties and mental illnesses
- the determinants of mental health problems and mental illnesses
- the performance of the mental health system overall.

The results of a recent evidence review indicate that culture is significantly and positively associated with physical health, social and emotional wellbeing, and reduces risk-taking behaviours[32]. Domains of culture include country, knowledges and beliefs, language, self-determination, family and kinship, and cultural expression. The review also provides examples of indicators[33] aligned with these domains, that evaluators could consider drawing on.

Resources

Ngaa-bi-nya[34] is a framework that offers a practical guide for the evaluation of **Aboriginal and Torres Strait Islander** health, social, and justice programs. Ngaa-bi-nya means to examine, try, and evaluate in the language of the Wiradjuri people of central NSW. It has been developed to reflect Aboriginal and Torres Strait Islander peoples' contexts and

[31] Dudgeon, P. et al (2016). Solutions that work: What the evidence and our people tell us: Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report. School of Indigenous Studies, University of Western Australia Centre for Child Health Research.

[32] Bourke, S., Wright, A., Guthrie, J., Russell, L., Dunbar, T. Lovett, R. (2018). Evidence review of Indigenous culture for health and wellbeing. The International Journal of Health, Wellness and Society. 8(4). https://www.researchgate.net/profile/Sarah-Bourke-2/publication/327225093_Evidence_Review_of_Indigenous_Culture_for_Health_and_Wellbeing/links/5d28b5e6a6fdcc2462da0f85/Evidence-Review-of-Indigenous-Culture-for-Health-and-Wellbeing.pdf

[33] Salmon, M., Doery, K., Dance, P., Chapman, J., Gilbert, R., Williams, R. & Lovett, R. (2019). Defining the Indefinable: Descriptors of Aboriginal and Torres Strait Islander Peoples' Cultures and Their Links to Health and Wellbeing, Aboriginal and Torres Strait Islander Health Team, Research School of Population Health, The Australian National University, Canberra. https://openresearch-repository.anu.edu.au/bitstream/1885/148406/8/Defining_the_Indefinable_WEB2_FINAL.pdf

[34] Yulang Indigenous Evaluation (2023). Ngaa-bi-nya. https://yulang.com.au/ngaa-bi-nya-evaluation-framework/

prompts the user to consider the historical, policy, and social landscape of Aboriginal and Torres Strait Islander people's lives. It provides a structure for the future development of culturally relevant, effective, translatable, and sustainable programs required for Aboriginal and Torres Strait Islander populations.

ADMINISTRATIVE DATA AND LINKED DATA

Administrative data routinely collected by health services is a key source of data to inform evaluation of the outcomes of mental health and suicide prevention programs and provide information at the systems level.

At the time of publication, these data sets can include the following:

- Medicare Benefits Schedule (MBS), which provides mental health items
- Pharmaceutical Benefits Scheme (PBS), which provides psychotropic medications
- Emergency Department presentations, including Core ICD-10 self-harm codes
- Community mental health care National Minimum Data Set
- Mental Health Establishments NGOs National Minimum Data Set
- Mental Health National Outcomes and Casemix Collection (NOCC)
- Seclusion and Restraint (SECREST)
- Activity Based Funding Mental Health Care Classification (ABF MHC) NBEDS
- Ambulance data
- Admitted patient data sets
- Suicide and self-harm monitoring projects
- Death and suicide data
- Residential Mental Health Care (among many others).

Administrative data can have value even when unlinked. For example, trends by age, region, and other dimensions will reflect the effectiveness of targeted programs. It can also inform process-related questions to understand the reach of a program and gaps in services.

Continued effort by governments has led to good consistency for many of the data collections listed above. These data sets are often hosted by the Australian Institute of Health and Welfare (AIHW) and based on contributions by various jurisdictions. This improves the ability to evaluate across different States and Territories in a consistent way. Even if this consistent data is used, it will generally be more timely to draw data from a jurisdiction directly.

Why use linked data: Administrative data is particularly powerful when linked. More routine data linkage is part of the National Agreement and can serve a range of purposes:

- Seeing longitudinal pathways for people receiving support through various programs.
- Improving statistical control (finding people who are otherwise similar to a program as a comparison group, or just a better understanding of risk factors).
- Testing impact on additional outcomes (e.g., mental health re-presentations) that would not otherwise be available.

- Tying shorter-term outcomes (e.g., K-10 improvement in a program) to longer-term ones. For example, participant-reported improvements can be linked to reduced rates of hospital admission or other adverse outcomes.
- Providing more of a 'systems' view – particularly relevant for systems evaluation, but important more generally for understanding how people are interacting with other services.

Accessing linked data almost always requires approval via a Human Research Ethics Committee, the approval and cooperation of relevant data custodians, and work by a linkage agency.

Considerations in deciding whether to use linked data: Some key considerations regarding the decision to use data linkage include the following:

- Linking program participants onto existing collections requires identifiers to be collected (such as name, date of birth, address, and Medicare number). The collection and permissions attached to identifiers will affect the feasibility of linkage.
- Custodian approval and data provision can be time-consuming, given strong governance processes. While processes will be more streamlined over time (for example, using enduring linkage programs such as the NDDA or PLIDA), governance processes are rigorous.
- Data lags similarly need consideration – 6 to 12 months delays in administrative data are not uncommon. Sometimes state or territory data provision (rather than national collections) can be more timely.
- Privacy requirements are typically tight. Small cohorts can have identifiability concerns, limiting what analysis is possible. This can affect analyses of less common events (such as numbers of people who die by suicide).
- Care is needed in interpretation of service use. In some cases, an increase in service use could be considered a good thing (e.g., a person managing a condition by regularly taking medication or having psychology appointments), and other times not (e.g., hospital admissions may indicate a deterioration).
- If using linkage to create a comparison group, a strong understanding of referral pathways is needed for a program to assess and manage the risk of selection effects that distort any comparison.
- There are inherent limitations to administrative data for assessing mental health outcomes. For instance, ongoing use of antidepressant medication (using PBS data) could sometimes be considered a good outcome (a condition is being managed well), a poor outcome (a condition has not resolved), or a spurious outcome (a drug is being used for a secondary purpose).

It places requirements on the quality of data collected by a program so that people can be linked, so is less suited to some forms of support services.

When it might be most appropriate: Using linked data is usually appropriate for larger scale evaluations to ensure sufficient time, budget, and a large enough sample size to identify findings in the linked data.

ADDITIONAL MEASURES

When considering additional quantitative data collection, the benefit and the burden of data collection should be balanced.

One consideration for additional data is to match data items to existing population data (for example that collected through the National Study of Mental Health and Wellbeing, HILDA). This can provide – albeit, a fairly weak – comparison between the people in the program and the general population.

One additional type of measure to consider is a quality of life measure. This could be useful for supporting economic evaluation, support synthesis or outcomes across programs, and enable comparisons of outcomes across health conditions.

Given the broader factors impacting mental health and suicide, considering these can also be useful in evaluation. Data on the social determinants of health could be useful to consider. The World Health Organisation lists the following as social determinants that can influence health equity in positive and negative ways:

- income and social protection
- education
- unemployment and job insecurity
- working life conditions
- food insecurity
- housing, basic amenities, and the environment
- early childhood development
- social inclusion and non-discrimination
- structural conflict
- access to affordable health services of decent quality.[35]

Data on social determinants of health

The Australian Institute of Health and Welfare identifies existing data on these social determinants.

What Australian Government data can be shared

From October 2022, All organisations can make a request to access Australian Government data through Dataplace.

Free, open datasets are available at data.gov.au. Open data includes the datasets published by Australian, state and local government agencies, as well as publicly-funded research data and datasets from private institutions that are in the public interest.

Share Data | Office of the National Data Commissioner

[35] World Health Organisation (2022). Social determinants of health. Geneva: WHO. https://apps.who.int/gb/ebwha/pdf_files/EB152/B152_22-en.pdf

7.2.4 ECONOMIC EVALUATION

This section identifies considerations for the use of different economic evaluation approaches depending on the context and available data. More detailed guidance on the technical aspects of evaluation is available from state and territory governments and other sources (see below).

THINKING ABOUT THE ECONOMIC BENEFITS OF MENTAL HEALTH AND SUICIDE PREVENTION PROGRAMS

Mental ill-health and suicide have significant social and economic costs – for individuals, their families and supporters, governments, and society. The Productivity Commission Inquiry into Mental Health (2020) recognised several broad categories of economic benefit resulting from improved mental health, including:

- reduced disability and death related to mental health conditions and suicide
- improved productivity (for example, higher employment rates, lower absenteeism and presenteeism)
- reduced government expenditure (for example, due to reduced use of emergency hospital services, reduced need for other health services).

This provides a scaffold for thinking about the benefits to be assessed in economic evaluations of mental health and suicide prevention programs. However, economic evaluation of mental health and suicide prevention programs is challenging, given the limitations of costs and outcomes data, barriers and timelines to access data, the difficulty of quantifying all benefits, and the complexity of factors affecting mental health and suicide.

It will not be possible or appropriate to try to put a dollar value on the benefits resulting from all mental health and suicide prevention programs, but it is possible to strengthen analysis of program costs – which can help to consider whether efficiencies might be possible – and to undertake economic evaluation when appropriate.

THE DATA NEEDED

The first thing required for this kind of analysis is complete data on costs/ expenditure. This should cover actual costs, rather than budgeted costs. For new programs, it should split start-up costs from ongoing costs. And for all programs, it should break down costs into consistent categories, such as staffing.

The second input required is complete and consistent quantitative data on outcomes for participants. This can include individual outcomes measures, such as those collected through standard outcomes data collections, as well as administrative data collected by health services on service usage.

For some economic evaluation methods, such as Cost-Benefit Analysis, a certain kind of outcomes measure – one that can be monetised – is required.

Economic evaluation also requires at least two alternatives to compare in terms of costs and outcomes – one of these options might be the current standard care or the "status quo".

KEY CONSIDERATIONS

A number of key considerations are important in designing effective value for money or economic evaluations. A brief overview of these is provided below, with more detailed guidance available in Australian Government and state and territory resources listed at the end of this section.

- **Comparison against a baseline or counterfactual**: Ideally, economic evaluations should compare costs and benefits of a policy intervention against a baseline – typically, what benefits and costs would have occurred if the intervention had not been implemented. This relies on experimental or quasi-experimental evaluation designs.
- **Who benefits and who incurs costs**: A large number of stakeholders may potentially benefit from a mental health and suicide prevention policy intervention. Evaluations should always make clear which stakeholders are included in the analysis of costs and benefits. Common categories of stakeholders who may benefit include funding agencies, broader government, service users, their friends, family and support persons, and the broader community. For example, an evaluation of a new mental health program may focus solely on costs and benefits for the program's clients and the agency that funds it, or it could consider additional impacts on other groups (such as benefits to the client's family, or reductions in the client's need for support from other areas of government). The approach used and rationale should always be made clear. Examples of costs and benefits that may impact people other than the service user and the funding agency (the 'referent groups') include:
    - **externalities/ spillovers:** Secondary benefits or costs that are not immediately apparent, often accruing to parties not in the groups directly considered. In cases where positive spillovers are significant, they can justify the investment in the program, for example, beneficial impacts on family members.
    - **flow-on effects:** The impact of spending generated by the program on the economy. These should be considered only in so far as they are real. That is, they should be included only if these effects would not have been generated by an alternative project or use of funds. In practice, these are typically not allowed for, as there is rarely evidence that spending on an alternative project would not generate similar flow-on effects.
- **Time horizon for measuring costs and benefits**: A suitable time horizon for measuring costs and benefits should be selected and made clear, considering the nature of the program, stage of implementation, and objectives of the evaluation.
- **Sensitivity analysis**: There may be considerable uncertainty in the assumptions used to estimate costs and benefits. These should be made transparent. Sensitivity (what-if) analysis should be performed to assess how changes in different assumptions would affect the conclusions drawn from the evaluation. A range of results should be reported allowing for variation in the assumptions.
- **Discounting future costs and benefits**: It is well established that individuals and governments would prefer to receive a dollar now, rather than a dollar in the future. This time preference should be taken into account by discounting costs and benefits that

occur in the future. Consideration is needed in selecting an appropriate 'discount rate' – see the resources at the end of this section.

- **Allowing for inflation**: Inflation is another reason that a dollar in the future is worth less than a dollar now. Costs and benefits should be shown consistently either in constant or current dollar values. Where constant (current) dollar values are used, real (nominal) discount rates should be applied.

Resources: State, territory and Commonwealth guidance on economic evaluations
State, territory and Commonwealth governments have produced a wide range of guidance materials with more details on economic evaluations. Examples include:

- NSW Treasury (2023). *Guidelines: Cost-Benefit Analysis*. NSW Government.
- Department of Finance (2021). *Commonwealth Evaluation Policy.* Australian Government.
- SA Department of Treasury and Finance (2014). *Guidelines for evaluation of public sector initiatives. Part B Investment Evaluation Process.* SA Government.
- Vic Department of Treasury and Finance (2023). *The Resource Management Framework.* Vic Government.
- Vic Department of Treasury and Finance (2023). *Economic Evaluation* Vic Government.
- Queensland Treasury (2020). *Program Evaluation Guidelines. Second Ed* Queensland Government.
- WA Department of Treasury (2020). *Program Evaluation Guide* Government of Western Australia.

COST ANALYSIS

For programs at an early stage of development, it may not be appropriate to conduct a more comprehensive economic evaluation, but it should be possible to assess the costs.

The costs of delivering a program often include:

- salaries for staff, e.g., support staff and management
- administration costs, e.g., IT, rent, and insurance
- other costs, e.g., staff training, travel, and vehicle expenses.

Where relevant to the program, it may also be important to consider out-of-pocket expenses incurred by people who access programs and services.

For new programs, start-up costs should be isolated from ongoing program costs to provide a more accurate picture of costs.

It may also be possible to compare the unit cost of delivering the program to the status quo or to a similar alternative program. This may be in the form of a cost per person supported or cost per hour of support provided. In doing so, the limitations of comparing only costs (without taking into account benefits) should always be made clear – for example, it is possible that a new service may cost more than the status quo, but still be value for money if it produces better outcomes. Equity should also be a key consideration, noting that it can cost more to reach certain populations.

Exploring costs – either the make-up of program costs or comparisons to other programs – can help to identify whether efficiencies are possible without compromising effectiveness. When looking at potential efficiencies, it is important to consider what costs could be streamlined without compromising outcomes, equity, or accessibility. For example, it might be more cost efficient to deliver a service over the phone, but if other data suggests the face-to-face relationship is critical to outcomes, this is unlikely to be a viable approach.

What data is required: This analysis requires information on the cost of delivering the program, which can often be found in the financial statements. To understand the unit costs of delivering the program, service activity data is needed in the form of number of persons supported or number of support hours provided.

COST-MINIMISATION ANALYSIS

Cost-minimisation analysis (CMA) is the simplest form of economic analysis. This type of analysis is suitable for situations where two or more programs have been demonstrated to be equivalent in terms of their health outcomes and the objective is to identify which of the equivalent programs is the least costly.

Cost-minimisation analysis has limited applicability (Table 6).

TABLE 6. COST MINIMISATION APPLICABILITY

Advantages	Disadvantages
• It is straightforward because it only compares costs. This means that the data burden is low. Data on costs are more straightforward to access relative to data on outcomes. • Useful when the effectiveness of two or more similar programs is established to be equivalent.	• Applicability is limited since it assumes mental health outcomes of the compared programs are equivalent. • It doesn't provide information about the value or benefit of an intervention in relation to its cost.

What data is required: This analysis requires two programs with outcomes data defined in the same terms. For a CMA to be valid, the assumption of equivalent health outcomes must be robust. Cost-minimisation is only appropriate when the difference in effects between programs is known to be small, and the cost difference is large. If the difference in effects is not known to be small from prior knowledge, the difference should ideally be established by an equivalence trial. If there's any uncertainty about the equivalency of outcomes, other methods of economic evaluation should be considered.

COST-EFFECTIVENESS ANALYSIS

Cost-Effectiveness Analysis (CEA) is an economic evaluation tool used to compare the relative costs and outcomes (effects) of different programs using a common outcomes measure to determine which offers the best value for money. Typically, it measures costs in monetary

terms, and outcomes in natural units, such as life-years gained, or symptom-free days (rather than in monetary terms).

Often, the focus of CEA is on the incremental cost-effectiveness ratio (ICER). This measure indicates the additional cost of achieving an additional unit of outcome with one intervention compared to another. The ICER is more sophisticated than a simple average cost-effectiveness ratio (ACER) because it helps decision-makers understand the specific value added by the new intervention, taking into account the existing standard of care, providing better information for resource allocation and treatment strategies.

While it has advantages and disadvantages (see Table 7 below), it seems to be more commonly used in mental health evaluation. [36]

TABLE 7. COST EFFECTIVENESS APPLICABILITY

Advantages	Disadvantages
• Provides policymakers with information about how to achieve the best outcome for a given budget or the lowest cost for a desired outcome. • By not converting mental health outcomes into monetary units, it avoids the ethical issues related to placing a monetary value on health and wellbeing, and concerns about the accuracy of the calculations. • Different types of outcomes can be used depending on the program, such as symptom-free days for depression, or life-years gained for interventions that increase lifespan.	• If different interventions (or studies) use different outcomes, it becomes challenging to compare them directly. • CEA often requires assumptions about future costs and outcomes, which can introduce uncertainty. • CEA might not be able to capture long-term program benefits. • Some broader benefits, especially those not directly related to mental health, might be excluded. For example, productivity increments post an intervention. • May not capture consumer defined recovery.

What data is required: CEA requires consistent collection of outcomes data for participants.

CEA is normally possible with the minimum data set for the mental health and suicide prevention measures collected.

Linked data would generally not be required. The exception to this would be if there is a strong desire to understand an outcomes measure not available in the minimum data set, e.g., the reduction in time in hospital per dollar spent on a program.

[36] Le et al. (2021). Cost-effectiveness evidence of mental health prevention and promotion interventions: A systematic review of economic evaluations. PLoS Med. 2021 May 11;18(5):e1003606. https://doi.org/10.1371/journal.pmed.1003606

COST-BENEFIT ANALYSIS

Cost-Benefit Analysis (CBA) involves comparing the total costs to the benefits of a particular program or policy, with both the costs and benefits expressed in monetary terms. The main objective of CBA is to determine if an investment or decision is sound by verifying if the benefits outweigh the costs, and by how many dollars.

While CBA is often considered the "gold standard" in assessing the economic feasibility and desirability of programs, it has been less commonly used in mental health and suicide prevention programs, for which its disadvantages can be particularly constraining (See Table 8).

TABLE 8. COST-BENEFIT ANALYSIS APPLICABILITY

Advantages	Disadvantages
• By understanding which programs or policies offer the greatest net benefit, service providers and governments can ensure that resources (money, time, labour) are allocated to where they will generate the most value. This can help to rank alternatives when resources are limited. • CBA allows for the comparison of multiple programs on a consistent scale to inform decision-making. • The process of CBA often involves considering different scenarios, including best-case and worst-case projections. This helps in identifying potential risks and uncertainties associated with a program or policy. • The process of CBA requires considering long-term impacts, not just immediate costs and benefits, to ensure sustainable decision-making. • CBA encourages the consideration of externalities, or the indirect costs and benefits that affect third parties. This can lead to a more holistic understanding of the impact of a program or policy.	• Fluctuations in mental health and suicide rates can make benefits difficult to establish and attribute. • Some benefits are intangible or difficult to assign a monetary value to, especially in the mental health context. • The choice of discount rate or the method to assign monetary value to specific benefits can be subjective, leading to greatly different outcomes based on different assumptions. • Predictions of long-term costs and benefits can be of dubious value, especially in dynamic environments. • Focusing solely on monetary values may cause equity concerns to be overlooked. For example, a program or intervention might have a cost benefit but disproportionately impact a vulnerable segment of the population. • May need to make numerous assumptions.

What data is required: A CBA requires that all expenses related to a program or policy are identified and calculated. This includes not only the direct and indirect costs but also opportunity costs (loss of other alternatives when one alternative is chosen), and any other associated expenses. Likewise, it requires that all the benefits from a program or policy are identified, quantified and assigned a monetary value. This can be straightforward in some cases but challenging for others.

Since costs and benefits often occur over different timeframes, there is also a need to bring future values into present terms through a process called discounting.

A CBA would be very limited by using only the minimum data sets for mental health and suicide prevention programs, due to the limited number of outcomes being captured. It would require the collection of additional data on outcomes for individuals.

In many cases, the use of linked data is likely needed to capture the full range of benefits. For example, if a program was able to demonstrate a reduction in the length of hospital admission stay, a program benefit could be calculated related to the avoided days spent in hospital. The administrative data sets likely to be most useful for economic analysis are:

- Emergency Department presentations
- community mental health care
- ambulance data
- admitted patient data sets
- Medicare Benefits Schedule (MBS), which provides mental health items
- Pharmaceutical Benefits Scheme (PBS), which provides psychotropic medications.

COST-UTILITY ANALYSIS

Cost-Utility Analysis (CUA) is a subtype of cost-effectiveness analysis that compares the costs (in monetary terms) of different programs to their outcomes in terms of utility, typically to the patient. The outcomes in CUA are measured using scales such as quality-adjusted life years (QALYs), disability-adjusted life years (DALYs), healthy years equivalent (HYE), and wellbeing years (WELLBY).

Compared to CEA, it requires additional assumptions and statistical modelling to transform observed outcomes into utility terms. While it is compelling because it enables standardised comparisons, its application is challenging in mental health (see Table 9).

TABLE 9. COST UTILITY ANALYSIS APPLICABILITY

Advantages	Disadvantages
• Typically, the outcomes measures in CUA combine both the quality and quantity of life into a single metric. This allows for direct comparisons between diverse programs affecting different aspects of health. • By translating health outcomes into a standardised unit (like QALYs), policymakers can compare the cost-effectiveness of programs across various health domains to make informed decisions about allocating resources to maximise health benefits within a constrained budget.	• Accurately measuring and valuing outcomes measures can be challenging. For example, different tools or instruments used to measure QALYs might yield slightly different values, leading to variability in outcomes. • Some stakeholders think the use of QALY type outcomes measures are discriminatory, especially towards the elderly or those with disabilities. This is because the value assigned to a year of life in perfect health might differ across populations, potentially leading to

Advantages	Disadvantages
• Measures like QALYs are able to incorporate patient or societal preferences about different health states, making the outcomes more reflective of values and preferences of the population. • Outcomes measures in CUA consider both morbidity and mortality, offering a more comprehensive view of impact on mental health. • Service providers and the government can establish thresholds (e.g., cost per QALY) to guide decisions about what programs are considered cost-effective.	decisions that favour certain groups over others. • The utility values used to calculate QALYs often come from subjective measures of wellbeing or health state preferences, which can vary among individuals and cultures. • The results of CUA from one context or population might not be directly applicable to another. Differences in health care costs, population demographics, or health state preferences can influence outcomes. • CUA often requires advanced economic and statistical methods, requiring technical expertise. • The issue of whether and how to discount future health benefits (i.e., giving them less weight than immediate benefits) can be contentious and can significantly influence CUA results.

What data is required: This analysis requires outcomes measured using scales such as quality-adjusted life years (QALYs), disability-adjusted life years (DALYs), healthy years equivalent (HYE), and wellbeing years (WELLBY).

A CUA would not be possible using only the minimum data sets for mental health and suicide prevention programs. It would require the collection of additional data on outcomes for individuals. It is unlikely that linked data would be required.

DECIDING WHAT APPROACH TO USE

To consider what approach is most appropriate should involve considering the following:

- **The availability and quality of data**: The most important consideration is what data is available. The above sections provide guidance on what data is required for different types of economic analysis. Most of the methods will require data additional to standard minimum data sets and this needs to be planned for to allow the collection of robust data. CBA will require more comprehensive data than CEA.
- **Program scale**: A more comprehensive economic evaluation will be more warranted for larger scale programs with more significant investment. Geographic diversity leads to variations in health care practices, costs, and outcomes. CEA provides a consistent method to evaluate the cost-effectiveness of interventions across these diverse regions, ensuring that recommendations are tailored to local contexts while still maintaining a standardised comparison metric.
- **Maturity of the program**: As programs mature, their outcomes tend to stabilise and become more predictable.

- **Nature of the program**: Some programs offer benefits beyond just health outcomes, impacting areas like societal wellbeing or economic productivity. In such cases, CUA and CBA capture these varied benefits, providing a more comprehensive assessment.
- **Duration of the program**: Longer-programs have outcomes and costs that change over time. CEA and CBA can evaluate the program's cost-effectiveness over its entire duration, accounting for changes in health outcomes and costs.

7.2.5 META-ANALYSIS

One of the benefits of having a more consistent approach to evaluation, and sharing individual evaluation results, is the ability to provide a better understanding of the mental health and suicide prevention system. Together, if shared, individual evaluations will be a powerful source of evidence for driving more consistent and effective responses to improve mental health and wellbeing and reduce suicide.

A meta-analysis may seek to identify patterns and themes from individual evaluations to generate findings that may be useful for developing new activities or refining existing activities. Meta-analysis takes the data from evaluations of the same type of program to determine overall patterns about how and for whom outcomes were generated.

Where appropriate and relevant, meta-analysis could involve:

1. classifying the evaluations by type of initiative.
2. the development of a framework to guide the extraction and re-analysis of data from each evaluation.
3. a review of each evaluation, extraction and analysis of quantitative and qualitative data.
4. a synthesis of findings and preparation of a brief report, or chapter in the final evaluation report about lessons learnt from different types of programs.

A modified version of the EMMIE framework[37] may provide a useful starting point for sharing information to continue building an evidence base of what works in varying contexts. This framework was developed for policymakers seeking practical evidence for decision-making about social policy and crime prevention programs to provide a consistent means of reviewing and extracting information from different types of evaluations.

Using EMMIE, the following information could be extracted from evaluations for sharing:

E the effectiveness or outcomes of an evaluation (where relevant this could include the effect direction and size of an outcome and the confidence that should be placed on that estimate)

M the means by which the outcomes were achieved (the mechanisms activated by the policy, practice or program in question)

M the moderators or contexts relevant to the evaluation (what caused the outcomes to vary for different people in different places)

I the insights into the success and failure in implementing the policy, practice or program, and system-level factors

E the economic costs (and benefits) associated with the policy, practice or program.

[37] Johnson, S., Tilley, N. & Bowers, K. (2015). Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews. Journal of Experimental Criminology. 11 459–473. https://doi.org/10.1007/s11292-015-9238-7

The Framework provides a template for sharing information in those circumstances where the sharing of full reports may not be possible, as well as to facilitate meta-analysis where full reports are shared (see Appendix 6).

7.3 ENGAGING SERVICE USERS IN EVALUATION

It is important that evaluations of mental health and suicide prevention programs engage people with lived and/or living experience who access, seek to access, or may have tried to access services (referred to throughout this section as service users), as they can give valuable insights about their experience, including which aspects of a program contribute to, or hinder, its effectiveness.[38]

Standard outcomes measures and administrative data alone are likely to be insufficient to capture the full value of a program to the people accessing it. Qualitative data can help to understand what matters to people accessing the program and its value to them, as well as help interpret the administrative data.

It can also help to understand the experiences of people who had trouble accessing the program or service, and who did not feel safe to do so although it can be difficult to reach these groups in evaluation.

TRAUMA-INFORMED APPROACH

Taking a **trauma-informed approach** to engaging people with lived and/or living experience in evaluation is essential to helping them feel comfortable and safe to contribute and to minimise the risk of their contribution resulting in traumatisation. A trauma-informed approach involves:

- prioritising safety – seeking to avoid re-traumatisation by not requiring someone to repeat events and stories, emphasising physical, psychological, and emotional safety
- understanding the neurological, cognitive, spiritual, biological, psychological, behavioural, and social impacts of trauma and violence, which at times can be lifelong
- recognising indicators that a person has experienced trauma
- responding appropriately through sensitive engagement and compassionate response protocols that emphasise choice and control.

Principles of trauma-informed care[39]
Key principles of trauma-informed care are safety, trustworthiness, choice, collaboration, and empowerment.

[38] Shand, F., Woodward, A., McGill, K., Larsen, M., Torok, M. et al. (2019). Suicide aftercare services: an Evidence Check rapid review brokered by the Sax Institute (www.saxinstitute.org.au) for the NSW Ministry of Health. https://www.saxinstitute.org.au/wp-content/uploads/2019_Suicide-Aftercare-Services-Report.pdf

[39] Blue Knot Foundation (n.d.). Talking About Trauma - Guide to Everyday Conversations for the General Public. https://blueknot.org.au/product/talking-about-trauma-guide-to-everyday-conversations-for-the-general-public-digital-download/

This means that evaluators need to:

- recognise the impact of power, and aim to maximise autonomy, choice, and safety
- recognise that response to trauma varies widely between individuals
- understand that the behaviours of people affected by trauma can represent adaptive responses to past traumatic experiences
- understand how trauma can impact people's relationships with others, including with services
- understand that the impacts of trauma can be cumulative across the lifespan of the individual
- understand the specific ways that trauma acts on/ impacts children – their social, emotional, and cognitive development, and attachment
- identify the importance of self-care for professionals and recognise the risks of vicarious trauma to workers and staff
- remain collaborative, consistent, truthful, and compassionate
- consider the physical and emotional safety of the person.

In practice, taking a trauma-informed approach applies to both design and data collection, and requires measures to be put in place to protect the wellbeing of both people with lived and/or living experience and staff. A trauma-informed approach will help to ensure processes are safe and accessible to people from priority populations.

SELECTING APPROPRIATE METHODS

The types of methods (and modes) for collecting qualitative data will depend on the context, type of information you are seeking, preferences/ needs of participants, and budget for the evaluation.

- **Individual interviews** are appropriate when the subject is sensitive and it may be uncomfortable for people to share in a group setting, you are interested in individual experiences with a service or pathways through a system, and/or you are intending to develop case studies to illustrate particular stories.
- **Group interviews** are appropriate when the subject matter is less sensitive, individual stories are not the main focus, and/or group experiences are of interest (e.g., peer support groups).
- **Observations** may be appropriate for group settings (e.g., peer support groups, co-design workshops) to gain insight into session characteristics, environment, presentation content, participant engagement and understanding, group dynamics, and overall observations. However, there might be sensitivities about using this approach.
- **Open text questions in surveys** are useful for allowing people to explain their responses to closed questions and describe their experiences and outcomes of a service in their own words. However, questions should be focused on the respondent's experiences and outcomes from the service to avoid the risk of distress and re-traumatisation.
- **Online forums** (i.e. a webpage where participants respond to prompts, questions and material, e.g., 'things I learned') allow people to participate at times that suit them without being publicly identified (using a pseudonym linked to a basic set of

demographic information that they can complete, and contact information that will only be visible to the evaluation team). This means they can provide feedback without the pressure or anxiety that could, for some, be associated with engaging with an unknown stranger.

- **Other asynchronous methods** (such as audio or video recordings, diaries, written documents) are useful for capturing longitudinal data and provide people with additional options for expressing themselves and their experiences. Asynchronous interviews, such as interviews conducted via text or web chat over several days, also give people time to think about and formulate their responses, which may be preferable or more accessible for some people.

SERVICE USER SELECTION

Consistent with the *National Statement on Ethical Conduct in Human Research*[40] (Section 8.1 and 8.2), before inviting service users to participate in the evaluation it is important to develop **inclusion and exclusion criteria** designed to minimise any potential risk to participants and ensure people with the most relevant experience are able to contribute. Some things evaluators may consider include:

- age of the participant
- length of time the participant has been accessing the program
- whether the participant is a current or past service user
- whether the participant has made an informed decision about their readiness to participate in the evaluation
- if there is increased concern about risk.

If people are able to consent to health or social care services, this can be considered an indication that they are able to consent to participate in an evaluation. People with mental health conditions and experiencing suicidal thoughts and behaviours are still generally able to participate in an informed consent process and have the right to make their own decisions about their readiness and willingness to participate in an evaluation. This is not for service staff to decide for them.

In the case of children and young people, there is a need to consider consent of a parent, carer or guardian in addition to the assent of the child.

If, during their participation, a service user becomes distressed (e.g., because sharing their experience brings up unexpected emotions, there are other things in their life they are dealing with on that day, they realise they misunderstood the focus of the evaluation, etc.), the evaluation team should be experienced at recognising these signs and have appropriate response protocols in place to manage this (see 'Data collection' section below).

[40] The National Health and Medical Research Council, Australian Research Council and Universities Australia (2023). National Statement on Ethical Conduct in Human Research. Canberra: National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2023

GATHERING PARTICIPANT INFORMATION AND INFORMED CONSENT

Evaluators should develop an **information sheet and consent form** that uses simple language to clearly outline the purpose of the evaluation, why the person has been invited to participate, and what participation will involve – including, the **benefits and potential risks** to the participant of being involved, **how their data will be used and stored**, and who to contact if they have questions. This will allow service users to make their own **informed decision** about whether to participate.

In designing the consent process, evaluators should consider their audience and produce information and consent processes tailored to them – for example, a video explaining the evaluation, or information sheets in different languages. When obtaining consent, it will also be necessary to collect other relevant information for the data collection, including:

- personal information – name, pronouns, contact details, state or territory (if relevant)
- communication preferences
- options for participation – e.g., mode for the survey, mode for the interview, preferred characteristics of the interviewer
- accessibility requirements
- preferred days and times for the interview
- anything else they would like the evaluation team to know.

If this information is collected online, the evaluation team could consider **calling or emailing the participant** (depending on their communication preferences) to confirm they have received their consent form and information, ensure they understand the purpose and nature of their involvement, and confirm a time and day for the interview. As this is the participant's first contact with the team, this call can help to introduce the team, alleviate any hesitation or confusion, and ensure the participant is comfortable to be involved. If the person does not answer, send a follow-up email with this information and the offer of a call back at a time that suits them.

If the person doing the interview is different to the person managing recruitment, it is also advised that **the interviewer contact the participant in the week before the interview** to introduce themselves, confirm the time for the interview, and answer any final questions they may have.

Interview questions should also be provided to participants ahead of time to allow them to prepare and reflect and identify if there are any questions they do not wish to answer. The evaluation team should make themselves available ahead of the interview to discuss with participants any questions they may not wish to answer.

Avoiding coercion

The information sheet should make it clear that:

- participation is **voluntary**
- their contributions will remain **anonymous**
- participation (or not) will not affect their **access to services and supports**, nor the **quality** of their support

- there are **no consequences** if they decide not to take part and they can choose **to opt out later** if they change their mind.

While services distributing invitations to participate in the evaluation is important for engagement, service users should be required to **contact the evaluation team to opt in**. This will avoid service users feeling pressure from the service and allow them to make their own informed decision about whether they would like to be involved. The evaluation team should provide service staff with advice about how to answer questions in order to help ensure service users do not feel pressured to participate.

REMUNERATION

Where possible, it is important to remunerate participants for their involvement in the evaluation to **ensure their time and expertise is recognised, respected and valued**. They should be provided with options for how they would like to receive this payment – cash, direct debit, gift card – to ensure they are paid in ways that are most valuable to them and avoid them feeling as though their participation has been tokenistic or a box ticking exercise.

Guidance on appropriate renumeration
Guidance on appropriate renumeration can be found in the *National Mental Health Commission's Paid Participation Policy 2020*

DATA COLLECTION DESIGN

Once eligibility has been determined, the next step is ensuring data collection tools are designed to minimise and manage the potential risk of distress to participants. Some ways to do this through the process of data collection involve the design of surveys or interviews/ group interviews are listed below.

- Survey
  - questions are focused on the person's experience with the service
  - questions are mostly selection-responses rather than open text
  - all questions are voluntary, with the exception of those that are necessary for determining which sets of questions people are asked
  - the format of the survey is accessible (e.g., consider text size, questions in audio format, images to support the text)
  - participants are reminded in the survey introduction that they do not need to answer any questions if they do not feel comfortable or would prefer not to
  - the survey introduction includes a link to a readiness to contribute guide
  - participants are provided with mental health support line numbers and other relevant contacts before the survey and once the survey is complete.
- Interviews/ group interviews
  - guides focus on experience with services, rather than the underlying factors and situation that contributed to mental ill-health or suicidality
  - guides are informed by an understanding of trauma-informed practice
  - guides are designed to provide participants with the opportunity to tell their story, following the flow of their engagement with a service

- questions are designed to allow participants to answer in a level of detail they determine
- some narrative text is included around the questions, so there is a clear rationale for what is being asked and to give participants some breathing space between questions
- examples are included as prompts in case participants are unsure of the meaning of the question.

Where possible, involving people with lived and/or living experience of mental ill-health and/or suicide, and those with experiences as families, carers and support people, in design can assist to identify questions that may be distressing or cause harm, and help to **ensure language is accessible and appropriate** for service users participating in the evaluation (see Appendix 3 for detailed guidance about involving people with lived and/or living experience of mental ill-health and/or suicide as part of the evaluation team).

Example: Methods for engaging with children and young people

Interviews with **children** may involve activities such as:

- 'emotion faces' for participants to describe the feeling associated with their response to a question
- sets of cards/ photos where children are asked to choose a card that describes what they think about their worker(s)
- asking children to draw their experience with the service using pen and paper, and then talk about or explain their drawing.

Interviews with **adolescents and young adults** may involve allowing them to choose between:

- an online forum
- asynchronous interviews
- individual interviews, with a greater focus on asking questions rather than using the activities.

OPTIONS FOR PARTICIPATION

It is also important to provide options for how service users can be involved so they feel safe and comfortable and their individual communication needs and preferences are addressed.

This will likely mean offering people choices between different methods for providing feedback (e.g., survey or interview) and options within each method.

- Choices for surveys
  - offer participants the choice between an online and paper survey
  - offer participants the option to have someone assist them to complete the survey if they wish.
  - Choices for interviews/ group interviews offer participants a choice about the mode of interview or group interview. This may be in person, via telephone, or via videoconference; the options you offer will likely depend on the available resources

of the evaluation and location of participants. If in person, interviews should be undertaken in a neutral, private and safe location such as the service offices, a park or other public location. Interviews should not be conducted in people's homes to maintain appropriate safety protocols and boundaries.

- For interviews, ask participants to request times that work for them. For group interviews, offer multiple groups on different days and at different times of day so participants can choose one that works for them
- offer participants the option of bringing along a support person of their choice if they wish
- where possible, offer participants some choice in terms of characteristics of the interviewer – a person with lived and/or living experience or not, gender, age (e.g., for young people), location (e.g., for rural and remote communities) and cultural and/or sexual identity.

There is also a need to consider accessibility, such as whether the participant will require language translating and interpreting services, including AUSLAN interpreters, and assistive communication devices for people with sensory or vision needs.

DATA COLLECTION

The following are some ways to ensure all stages of the interview are trauma-informed.

Start of the interview

- Introduce yourself and ask the participant how they are going.
- Check it is still a good time for the interview and whether they would still like to be involved.
- Thank them for making the time to speak to you and tell them a bit more about yourself if you wish.
- Remind them of how long the interview will take.
- Remind them of the purpose of the evaluation and what the interview will cover.
- Acknowledge that sharing one's personal experiences can be difficult and may bring up unexpected things either at the time or at a later date. Let them know that if this happens, they can take a break, stop the interview, or choose not to answer particular questions.
- Check they are somewhere they feel comfortable to have the discussion.
- Advise them there are no right or wrong answers, you want to know what they really think.
- Remind them how the information they provide will be used in reporting and that they will not be identified in any way, to reinforce that the information they provide is confidential.
- Explain that you have a duty of care to share what they have said if you are seriously concerned about their or someone else's safety, but that this does not often happen and that you will tell them if you are going to do this. You should also advise them who you will be telling if you have to do this.
- Ask if they have any questions before you begin.
- If recording the interview, explain the purpose for doing so (e.g., for notetaking purposes), and ask them if they are okay with this. If not, take notes.

During the interview

- Check in with the participant at least once during the interview to make sure they are feeling alright and ask whether they would like a break.
- Be attuned to the difference between someone reflecting emotionally on what was a difficult experience versus becoming distressed. Signs of distress may be overt – for example, shutting down, becoming quiet and withdrawn, dissociating – or more subtle, such as looking away. But it is important not to make assumptions about the reasons for an individual's behaviour and to check in.
- If the person becomes distressed, pause the interview and check in with the person. Assure them that it is completely okay to:
  - take a break
  - not answer a particular question or group of questions
  - pause the interview to resume at another time
  - or end the interview all together – if this is the case, let them know that none of the information they have provided in the interview will be used, and that they will still be remunerated for their time.
- Respect the participant's decision about whether or not to continue the interview. If they are upset but wish to continue, adjust the pace of the interview. If appropriate, pause and allow the interviewee space to re-ground in the moment by sensing the space around them, or by reflecting on a question, such as "can you tell me about all of the sounds you can hear at the moment where you are?"
- Remind the person about mental health support line numbers and provide those if necessary – acknowledging that some people may find these impersonal – and suggest the person contact their nominated support person.
- If the person is still accessing the service they are providing feedback about, or they are a past service user, offer to reconnect the person to the service for support, provided they have not just raised any issues about the service or their worker that would make this an inappropriate offer. If the latter is the case, refer the person to mental health support lines.

End of the interview

- Thank the participant again for their time, openness and honesty, and willingness to participate.
- Check in again about how they are feeling now the interview is over.
- Remind them to practice self-care in whatever form that takes for them (e.g., do something they enjoy, like catching up with a friend or going for a walk) over the coming days.
- If concerned, suggest they contact someone they would normally go to for support (e.g., a friend/ family member, GP, other clinical support) and remind them of the mental health support line numbers. For participants still accessing the service they are providing feedback about, or who have accessed the service in the past, offer to reconnect them to the service for further support.
- Advise the participant on next steps for analysis and reporting (be clear about whether or not the report will be published) and remind them about confidentiality. If the report is to be published, ask if they would like to be sent a copy.

- Advise them that they can contact you at any point if there is anything they wish to add or retract.
- Let them know you will send a follow-up text or email, depending on their preference, the next day.
- Provide follow up check-in options for the participant if necessary (see 'after the interview').

After the interview

- Send a follow-up text or email the next day to:
    - thank them for their contribution and explain how feedback will be considered/ what the next steps in the process might be
    - acknowledge that these experiences can cause discomfort or distress and encourage people to consider self-care and remind them about the mental health support line numbers
    - advise them of the amount they will be paid for participating, how this will be paid, and when they can expect to receive the payment.
- You might also consider offering the person the opportunity to review their transcript if they would like.

8. ENSURING THE EVALUATION IS CONDUCTED ETHICALLY

All evaluation must be conducted ethically – whether or not you need formal approval of a Human Research Ethics Committee (HREC). The ethical issues that may arise during the conduct of an evaluation include protection of privacy and confidentiality, risk of possible distress to participants, and sharing or distribution of benefits.

8.1 KEY ETHICAL CONSIDERATIONS

The *National Statement on Ethical Conduct in Human Research (2023)* outlines how to ensure evaluations are conducted ethically.

Key considerations for the ethical conduct of evaluations include the following components.

- **Informed consent**: People participating in an evaluation need to be aware of the purpose of the project, who or what group is funding it, how the findings will be used, if there are any potential adverse impacts of their participation, and who will have access to the findings. The main purpose of informed consent is to ensure that the participant has all the information they need to be able to make an informed decision as to whether they will participate in the evaluation or not. Additional information should also be provided in the event that the participant becomes distressed in any way during their participation.

- **Voluntary participation**: This means that people participate in the evaluation free from coercion. Participants are free to not participate in the evaluation at all, or to withdraw their participation at any time without negatively impacting on their involvement in the program (or future programs). It is the right of participants to leave a program of this nature at any time, therefore no pressure should be placed on those who choose not to continue. Explanations are also not required.

- **Do no harm**: Harm can be both physical and/or psychological and, therefore, can be in the form of stress, pain, anxiety, diminishing self-esteem, or an invasion of privacy. It is very important that the evaluation process does not in any way harm (unintended or otherwise) participants.

- **Confidentiality**: This means that any identifying information is not made available to, or accessed by, anyone but the evaluators. Confidentiality also ensures such identifying information is excluded from any reports or published documents. Given that there may be small numbers of participants in some programs, it is very important to consider how reports are worded to ensure that there is no opportunity for people to be identified even though names are not used.

8.2 FORMAL ETHICS APPROVAL

The distinction between research and evaluation is somewhat blurred and it can be difficult to judge whether formal ethics approval is required. Triggers for considering formal ethical review include an intention to use the data for secondary purposes (such as publication), to compare groups of people, or to test non-standard or innovative protocols or models of care[41]. The use of linked administrative data will almost always require formal ethical review. When Aboriginal and Torres Strait Islander people are represented among people who access the programs and services being evaluated at a higher proportion than exists in the population as a whole, there is a greater likelihood of the need for ethics approval[42,43]. Evaluations of programs for Aboriginal and Torres Strait Islander communities and those involving analysis of outcomes for Aboriginal and Torres Strait Islander communities will also have additional ethical considerations (see section 8.3).

The *National Statement on Ethical Conduct in Human Research 2023* (the National Statement) states that formal ethics approval by a Human Research Ethics Committee is required when research exceeds a 'lower risk' rating[44]. These ratings are given by the extent of the potential *harm* or *discomfort* caused by the research (Table 10). Organisations may choose to review research under alternative processes if it does not exceed a lower risk rating.

TABLE 10. ASSESSING ETHICAL RISK

Risk	Sub-risk	Description
Lower risk	Minimal	No risk of harm or discomfort; potential for minor burden or inconvenience
	Low	No risk of harm; risk of discomfort (with acknowledgement of potential burden)
Higher risk	Greater than low	Risk of harm (with acknowledgement of potential burden)
	High	Risk of significant harm (with acknowledgement of potential burden)

[41] National Health and Medical Research Council (2014). Ethical considerations in quality assurance and evaluation activities. https://www.nhmrc.gov.au/about-us/resources/ethical-considerations-quality-assurance-and-evaluation-activities#block-views-block-file-attachments-content-block-1

[42] Aboriginal Health and Medical & Medical Research Council Ethics Unit (2023). NSW Aboriginal Health Ethics Guidelines: Key Principles. https://www.ahmrc.org.au/wp-content/uploads/2021/02/AHMRC_Health-Ethics-guidelines-2023_01.pdf

[43]Australian Institute of Aboriginal and Torres Strait Islander Studies (2023). Ethical research. https://aiatsis.gov.au/research/ethical-research

[44] The National Health and Medical Research Council, Australian Research Council and Universities Australia (2023). National Statement on Ethical Conduct in Human Research. Canberra: National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2023

8.3 ETHICAL REQUIREMENTS FOR EVALUATIONS INVOLVING ABORIGINAL AND TORRES STRAIT ISLANDER COMMUNITIES

Evaluations of Aboriginal and Torres Strait Islander-specific programs and mainstream programs with analysis of outcomes for Aboriginal and Torres Strait Islander communities have specific ethics approval requirements.

Evaluations must demonstrate the core values of responsibility, reciprocity, respect, equality, survival and protection, and spirit and integrity; together with core principles related to consent, research agreements, cultural and intellectual property and cultural competency.[45,46,47]

A culturally safe evaluation:

- recognises the diversity and uniqueness of peoples and individuals
- recognises Aboriginal rights to self-determination and rights to control of research about their communities
- respects Aboriginal knowledge, practices, and innovations
- establishes relationships through consultation and negotiation, and recognises that consultation is an ongoing process and should achieve mutual understanding
- ensures free, prior and informed consent for research participants
- enacts Aboriginal people's right to full participation appropriate to their skills and experiences
- ensures research be of benefit to the community and reflects the needs and interests of the community
- shares the research findings with the researched community.

Ethical evaluation references for Aboriginal and Torres Strait Islander communities

Australian Institute of Aboriginal and Torres Strait Islander Studies (2020). *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*.

National Health and Medical Research Council (2018). Ethical conduct in research with Aboriginal and Torres Strait Islander Peoples and communities: Guidelines for researchers and stakeholders, Canberra: Commonwealth of Australia.

Australian Evaluation Society, 2021, *First Nations Cultural Safety Framework*.

[45] Australian Institute of Aboriginal and Torres Strait Islander Studies (2020). AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research. https://aiatsis.gov.au/sites/default/files/2022-02/aiatsis-code-ethics-jan22.pdf

[46] Commonwealth of Australia, Department of the Prime Minister and Cabinet (2018). Indigenous Advancement Strategy Evaluation Framework. https://pmc.gov.au/sites/default/files/publications/ias-evaluation-framework.pdf

[47] National Health and Medical Research Council (2018). *Ethical conduct in research with Aboriginal and Torres Strait Islander Peoples and communities: Guidelines for researchers and stakeholders.* Canberra: Commonwealth of Australia, p. 13. https://www.nhmrc.gov.au/about-us/resources/ethical-conduct-research-aboriginal-and-torres-strait-islander-peoples-and-communities.

INDIGENOUS DATA SOVEREIGNTY

It is important that evaluations consider the evolving requirements for Indigenous Data Sovereignty. Indigenous Data refers to information or knowledge, in any format or medium, which is about and may affect Indigenous peoples both collectively and individually.

An Australian Public Service-wide framework for Indigenous Data and Governance is being developed with Aboriginal and Torres Strait Islander partners. The Framework for Indigenous Data and Governance aims to improve the accessibility, relevance, interpretability, and timeliness of government-held data for Aboriginal and Torres Strait Islander peoples. The Framework for Indigenous Data and Governance will also explore the practical intersection between Australian Government objectives and those of the Indigenous Data Sovereignty movement in Australia. This work is aligned with Closing the Gap Priority Reforms Three and Four.

The Maiam nayri Wingara Indigenous Data Sovereignty Principles[48] assert that in Australia, Aboriginal and Torres Strait Islander people have the right to:

- exercise control of the data ecosystem, including creation, development, stewardship, analysis, dissemination, and infrastructure.
- data that are contextual and disaggregated (available and accessible at individual, community and First Nations levels).
- data that are relevant and empower sustainable self-determination and effective self-governance.
- data structures that are accountable to Indigenous peoples and First Nations.
- data that are protective and respect individual and collective interests.

Indigenous Data Governance enables Aboriginal and Torres Strait Islander people, and their representative and governing bodies, to accurately reflect Aboriginal and Torres Strait Islander stories, providing the necessary tools to identify what works, what does not, and why. Effective Aboriginal and Torres Strait Indigenous Data Governance empowers Indigenous peoples to make the best decisions to support their communities in the ways that meet their development needs and aspirations. [49]

[48] Maiam nayri Wingara. (2018). Indigenous Data Sovereignty. Communique. Indigenous Data Sovereignty Summit, 20th June 2018, Canberra, ACT. https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/63ed934fe861fa061ebb9202/1676514134724/Communique-Indigenous-Data-Sovereignty-Summit.pdf

[49] Maiam nayri Wingara. (2018). Indigenous Data Sovereignty. Communique. Indigenous Data Sovereignty Summit, 20th June 2018, Canberra, ACT. https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/63ed934fe861fa061ebb9202/1676514134724/Communique-Indigenous-Data-Sovereignty-Summit.pdf

9. ANALYSING AND REPORTING

9.1 ANALYSIS

Analysing data to summarise it and look for patterns is an important part of every evaluation. The type of analysis you use will depend on your evaluation design, the nature of the data collected and data quality. The tables below (Tables 11 and 12) describe some common types of analysis for qualitative and quantitative data.

TABLE 11. COMMON TYPES OF ANALYSIS FOR QUALITATIVE DATA

Type of analysis	Description of approach	Common applications
Thematic analysis	Iterative: Interviews, observation, documents, or group data are explored with an eye to identifying key emergent themes	Thematic analysis of this nature takes more time than other forms of analysis
	Prescriptive theory driven: Interviews, documents, or group data is mined for evidence that pertains to particular themes that have been identified, before analysis begins	Requires a strong theory or principles identified ahead of time that can then be applied to the data gathered
Narrative account	Various sources are used to build a chronological and analytical description of how events occurred over time	Describing the development of a program, describing how consumers move through a program
Illustrative	Excerpts or brief summaries from qualitative data (either quotations or descriptions) are used to illustrate the findings from quantitative data	When reporting is intended to provide users with the overall patterns and some understanding of what this looks like for individuals
Content analysis	Quasi-statistical Frequency counts of particular comments or words used	In highly structured qualitative evaluations, such as analysing media coverage of an event
Qualitative Comparative Analysis	Mixed methods approach Causal analysis of different combinations of factors that lead to outcomes	When the program is being implemented at multiple sites and it is important to understand the role of context to support transferability

TABLE 12. COMMON TYPES OF ANALYSIS FOR QUANTITATIVE DATA

Type of analysis	Description of approach	Common applications
Descriptive	Describing the data from a sample or population	Inferential – inferring the characteristics of a population from a sample of data
Univariate – one variable at a time	Frequency tables Mean, median, mode Range, standard deviation Graphical analysis	Chi-square goodness of fit T-test
Bivariate analysing two variables	Simple cross-tabulations Correlation co-efficient	Chi-square
Multivariate – several variables at once	Disaggregated cross-tabulations Multiple regression, logistic regression, structural equation modelling Factor analysis, cluster analysis	ANOVA, MANOVA

It can be helpful to do a trial run of the data analysis process to ensure that the data you are gathering addresses the key questions and information needs of key stakeholders.

It is critical that the analysis identifies the **limitations**. That is, any factors that may have affected the results either directly or indirectly.

Other resources on analysis

- *Better Evaluation: Rainbow Framework*
- *Commonwealth Evaluation Toolkit*
- *AIFS: Making the most out of evaluation*

9.1.1 SYNTHESIS AND INTERPRETATION

The next step is to synthesise and **interpret the data**. This involves bringing all data sources together to understand what – in combination – the data say about the program being evaluated. The qualitative data should be able to help unpack and explain patterns in quantitative data.

You may also consider discussing these findings with key stakeholders to make sense of the findings from multiple perspectives. This can help to ensure findings are interpreted in context. Engaging stakeholders and people with lived and/or living experience of mental ill-

health and/or suicide in the interpretation of data can ensure the range of factors impacting on the data are considered.

9.2 KEEPING STAKEHOLDERS INFORMED OF PROGRESS

It is valuable to provide regular updates to key stakeholders during the evaluation to keep them up to date with emerging findings, and to help staff and program managers in considering the implications of the evaluation for their work.

9.3 REPORTING

An evaluation report should answer the key evaluation questions to make a clear and contextualised judgement of the value of the program. When considering reporting, it is helpful to:

- clarify the **purpose** of the report. Will your report be used for decision-making, for learning and understanding, or for strategic or political purposes
- clarify the key **audience** or audiences for the report. It may be possible to meet the needs of all audiences in one report, but often it can help to complement a report with presentations, written formats, electronic documents, video or audio formats.
- develop an **outline** for the report. Often this may be structured around your key evaluation questions. It might also be structured around the logic model of the program.

The report should generally include a series of **recommendations or implications**. These are usually most useful – strategic and feasible – if developed collaboratively with those responsible for the program being evaluated.

All reports should make clear the **methods** used and their **limitations**, to facilitate interpretation of findings, and support future evaluation.

Reports should also acknowledge those who contributed to the report, such as staff members from the commissioning organisation, service users (though not by name), and evaluation team members, including those with lived and/or living experience.

See Appendix 6 for an example of a template, based on a modified version of the EMMIE framework[50] which may provide guidance for consistently sharing information about evaluation to continue building an evidence base of what works in varying contexts. This could be used together with the sharing guidelines.

Guidance on using evaluation findings and reporting can be found here:

- *Commonwealth Evaluation Toolkit*
- *Better Evaluation: Using Evaluation Findings*

[50] Johnson, S., Tilley, N. & Bowers, K. (2015). Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews. Journal of Experimental Criminology. 11 459–473. https://doi.org/10.1007/s11292-015-9238-7

APPENDIX 1 KEY CONCEPTS

This section will include definitions of key terms, such as monitoring, evaluation, outcomes, and impact.

Term	Definition
Access	People can obtain health care at the right place and right time, taking account of different population needs and the affordability of care.
Appropriateness	In evaluation, appropriateness may relate to the extent to which a program addresses an identified need or is suitable for a community or context. For example, the program is person-centered, culturally appropriate, rights-based, trauma-informed, and recovery oriented. Mental health consumers and carers are treated with dignity and confidentiality and encouraged to participate in choices related to their care. Consumers and carers report positive experiences.
Continuity of care	Ability to provide uninterrupted and integrated care or service across program, practitioners over time. Coordination mechanisms work for mental health consumers, carers, and health care providers. Care and support is holistic, and includes psychosocial and physical dimensions.
Cultural appropriateness	The extent to which evaluation activities recognise, respect, and respond to the culture and values of the participants and communities in which the program operates.
Effectiveness	The program achieves the desired outcomes from the perspective of the mental health consumer, support person or carer, and the clinician or peer-worker.
Efficiency and sustainability	The right care is delivered at minimum cost, and human and physical capital and technology are maintained and renewed *while* innovation occurs to improve efficiency and respond to emerging needs. Members of the workforce receive appropriate support and report positive experiences.
Equity	All people should have a fair opportunity to attain their full health potential, and no-one should be disadvantaged in achieving this potential if it can be avoided.
Evaluation	A process that can be used to systematically and transparently assess the merit or worth of a program.
Evaluation approach or design	The lens through which an evaluation will be viewed. Often a reflection of the purpose.
Evaluation methods	The tools (e.g., surveys, group interviews) and procedures (e.g., recruitment, consent) that will be used to collect and analyse data.
Indigenous Data	Information or knowledge, in any format or medium, which is about and may affect Indigenous peoples both collectively and individually.

Indigenous Data Governance	The right of Indigenous peoples to autonomously decide what, how, and why Indigenous Data are collected, accessed and used. It ensures that data on or about Indigenous peoples reflects their priorities, values, cultures, worldviews, and diversity.
Indigenous Data Sovereignty	The right of Indigenous peoples to exercise ownership over Indigenous Data. Ownership of data can be expressed through the creation, collection, access, analysis, interpretation, management, dissemination, and reuse of Indigenous Data.
Lived and/or living experience	People with lived and/or living experience of trauma, neurodiversity, mental health challenges, psychological distress, suicide, substance use or addiction, and their families, carers, supporters, those experiencing bereavement, advocates and allies.
Meta-analysis	Meta-analysis is the statistical combination of results from two or more separate studies. Potential advantages of meta-analyses include an improvement in precision, the ability to answer questions not posed by individual studies, and the opportunity to settle controversies arising from conflicting claims.
Monitoring	Monitoring is a process of collecting, analysing, and reporting information about the implementation and impacts of a program at regular intervals. Choose a few key indicators to track over time to indicate performance, support ongoing improvement, and inform evaluation.
Outcomes	A result that can be measured or observed in terms of a changed condition of people, organisations, or systems, attributable to a program.
Process evaluation	An evaluation of the implementation and delivery of a program. It focuses on the inputs, activities, and outputs of a program.
Program logic model	A one-page diagram that shows the components of a program and how they are expected to work together to deliver outputs and outcomes and contribute to an intended impact.
Service users	People with lived and/or living experience who access, seek to access, or may have tried to access mental health and/or suicide prevention services.

APPENDIX 2 OTHER RELEVANT DOCUMENTS

This Framework draws on, and intends to complement a range of other documents inclusive of but not limited to:

- The **Evaluation in the Commonwealth Toolkit** and state and territory governments evaluation guidelines, and the Department of Health and Aged Care's evaluation strategy.
- The **National Mental Health Performance Framework (NMHPF) 2020,** which was developed for facilitating a culture of continuous quality improvement in mental health service delivery. The Framework supports the Australian Government's and state and territory governments' commitment to improving accountability and transparency at the Mental Health Service Organisation level.
- The **Australian Health Performance Framework (AHPF),** which is a tool for reporting on the health of Australians, the performance of health care in Australia, and the Australian health system. The indicators within the Framework can be disaggregated and analysed through selected population groups, providing a rich source of information at the national, state and territory, and local levels (where data is available).
- The **Aboriginal and Torres Strait Islander Health Performance Framework,** which brings together information from numerous sources to provide a comprehensive, up-to-date view of the state of Aboriginal and Torres Strait Islander health outcomes, health system performance, and the broader determinants of health in one area. It is designed to inform policy, planning, program development, and research.
- The Productivity Commission's **Indigenous Evaluation Strategy,** which provides a whole-of-government framework for Australian Government agencies to use when selecting, planning, conducting, and using evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people. The Strategy puts Aboriginal and Torres Strait Islander people at its centre. It recognises the need to draw on the perspectives, priorities, and knowledges of Aboriginal and Torres Strait Islander people if outcomes are to be improved.
- **Closing the Gap** commitments and measures, which recognises the changes required in the way governments work with Aboriginal and Torres Strait Islander communities to close the gap, and tracks progress on key measures.
- The **National Strategic Framework for Aboriginal and Torres Strait Islander Peoples' Mental Health and Social and Emotional Wellbeing 2017–2023**. This Framework has a dedicated focus on Aboriginal and Torres Strait Islander people's social and emotional wellbeing and mental health. It sets out a comprehensive and culturally appropriate stepped care model that is equally applicable to both Indigenous-specific and mainstream health services.
- **Solutions that work: What the evidence and our people tell us, Aboriginal and Torres Strait Islander Suicide Prevention Evaluation Project Report,** which identifies success factors for Aboriginal and Torres Strait Islander Suicide Prevention, suggests a framework for evaluating these programs, and recommends the use of participatory action research for future suicide prevention research in Aboriginal and Torres Strait Islander communities.

APPENDIX 3 ENGAGING LIVED AND/OR LIVING EXPERIENCE TEAM MEMBERS

A3.1 KEY PRINCIPLES

Key principles that evaluators should follow when engaging with people with lived and/or living experience as team members in evaluation, include[51,52]:

- **Flexibility**: Team members with lived and/or living experience can choose how and when (e.g., at which stages of an evaluation) they want to be involved, balanced with the needs of the evaluation.
- **Accessibility and inclusion**: All stages of the evaluation are accessible to and inclusive of people with diverse LLE. LLE is understood and valued in terms of its diversity.
- **Communication:** Communication with team members with LLE is regular, accessible, and uses appropriate and accessible language.
- **Transparency:** Clear roles are defined and team members with LLE are informed about the scope of the engagement and boundaries of what is possible to achieve.
- **Supportive and empowering:** Power dynamics are considered and actively managed. Team members with LLE receive the emotional and practical support they may require to be successful and are supported to practise self-care.
- **Mutually beneficial:** Involvement is arguably as beneficial to people with LLE as it is to the organisation or evaluation itself. Team members with LLE are adequately remunerated for their time and expertise, given opportunities to further develop their skills, and work closely with evaluators to build lasting relationships and opportunities for future engagement.

> *"Ultimately, it is not solely a person's lived experience that matters, but the humanity and self brought to each interaction and experience and how they are supported to contribute to or lead the planning, design, implementation and evaluation of changes they seek in an ongoing way."*[53]

A3.2 READINESS

Engaging team members with LLE will work best when the right conditions are in place. Evaluation teams and funding organisations should consider their own readiness for working

[51] Maloney, J., Jackman, B., Marra-Brown, S., Dale, R., Lorigan, A., Elwyn, R. & Waltmon, J. (2023). Working with lived and living experience evaluators: A practical framework. Suicide Prevention Conference, Canberra.

[52] Jo Farmer (2018). Evaluating with a mental health lived experience. https://www.linkedin.com/pulse/evaluating-mental-health-lived-experience-joanna-farmer

[53] Hodges, E., Leditschke, A., Solonsch, L. (2023). The Lived Experience Governance Framework: Centring People, Identity and Human Rights for the Benefit of All. Prepared by LELAN (SA Lived Experience Leadership & Advocacy Network) for the National Mental Health Consumer and Carer Forum and the National PHN Mental Health Lived Experience Engagement Network. Mental Health Australia, Canberra. https://nmhccf.org.au/our-work/discussion-papers/the-lived-experience-governance-framework-centring-people-identity-and-human-rights-for-the-benefit-of-all

with team members with lived and/or living experience.[54] This includes aligning with existing organisational policies and frameworks that promote diversity and inclusion. Factors that contribute to organisational readiness include:

- leadership supportive of the engagement of people with lived and/or living experience and committed to centring their views, including leaders representing lived and/or living experience perspectives themselves
- open-mindedness to considering lived and/or living experience viewpoints, and to people with lived and/or living experience shaping evaluation in a meaningful way
- clear expectations about the role of people with lived and/or living experience that value the perspectives people with lived and/or living experience bring
- a commitment to learning and flexing over time if needed
- operational systems support collaboration (e.g., shared editing of documents)
- appropriate resourcing (including remuneration) for lived and/or living experience team members and for broader evaluation team members to take the time to effectively engage lived and/or living experience perspectives
- appropriate resourcing and training for evaluation team members
- training for broader evaluation team members, for example, to effectively manage any challenging thoughts that may emerge in mental health discussions
- adequate human resource strategies and resources to deal with any associated stigma or discriminatory practice that may emerge.

Teams can also provide resources to support people with lived and/or living experience to consider their preparedness to be part of an evaluation team. People with lived and/or living experience identified three key aspects of readiness people should consider when deciding whether to take on a lived and/or living experience role in evaluation:

- **emotional readiness**: anxiety or uncertainty about the process; mental and emotional capacity to take on the work; time since trauma, suicide attempt, or loss; supports and strategies in place to protect their own wellbeing; awareness of personal warning signs and triggers
- **skills and knowledge**: understanding lived and/or living experience of mental ill-health and suicide and the role of people with lived and/or living experience in evaluation (e.g., the purpose and benefits); knowledge of trauma-informed approaches and understanding how to navigate knowledge and skill gaps; understanding of the personal and unique nature of recovery; understanding of how to safely and appropriately use and share personal experiences
- **biases**: understanding the types of knowledge and expertise they prioritise and acknowledging their gaps.

It is important to remember that people with lived and/or living experience have the **dignity of risk** to be involved – that is, they have the right to assess the level of risk associated with the role and decide for themselves whether they are ready to take it on – they know themselves and their wellbeing best.

[54] Maloney, J., Jackman, B., Marra-Brown, S., Dale, R., Lorigan, A., Elwyn, R. & Waltmon, J. (2023). Working with lived and living experience evaluators: A practical framework. Suicide Prevention Conference, Canberra.

A3.3 STAGES OF ENGAGEMENT

When planning to engage with team members with lived and/or living experience, evaluators should consider how they recruit, support, and work with these team members, as well as how they involve them in the various project phases[55,56].

RECRUITMENT

People with lived and/or living experience can be recruited through the organisation running the program (e.g., engaged members, members of a lived and/or living experience advisory committee, former program participants), lived and/or living experience organisations, or broader mental health and suicide prevention organisations with pools of engaged members, or relevant job boards.

To support recruitment, it is important to develop an **overview of the role** that clearly and simply outlines:

- the evaluation
- the expected role of the lived and/or living experience team members
- the nature of lived and/or living experience required
- skills and experience required
- support and training that will be provided (for both wellbeing and skill building)
- remuneration.

Remuneration
Remuneration for people with lived and/or living experiences of mental ill-health and/or suicide, and those with experiences as families, carers and support people, is essential to ensuring their time and expertise is recognised, respected and valued. Remuneration for people with lived and/or living experience should be factored into the budget for an evaluation.

Guidance on appropriate renumeration can be found in the *National Mental Health Commission's Paid Participation Policy 2020.*

It is important to **ensure no candidates are disadvantaged by the assessment process** by making it accessible. This can include:

- providing candidates with the interview questions ahead of time to allow them to prepare

[55] Webb, M., Cooper, C., Hemming, L., Dalton, A., Unity, E., Simmons, M., Bendall, S. & Robinson, J. (2023). Guidelines for involving young people with lived and living experience of suicide in suicide research. Melbourne: Orygen. https://www.orygen.org.au/Training/Resources/Self-harm-and-suicide-prevention/Guidelines/Suicide-research

[56] Krysinska, K., Ozols, I., Ross, A. et al (2023). Active involvement of people with lived experience of suicide in suicide research: a Delphi consensus study. *BMC Psychiatry* 23, 496. https://doi.org/10.1186/s12888-023-04973-9

- options for how the interview is conducted to accommodate people's communication preferences (e.g., different modes for the interview, support to access these modes, the ability to respond to questions in writing).

The **number of people recruited** will depend on the size and nature of the evaluation, but it is considered best practice to recruit multiple lived and/or living experience team members to:

- account for people having different areas of interest in the evaluation
- account for people needing to take a break or withdraw from the evaluation due to conflicting priorities or changes to their mental health and/or suicidal distress levels
- ensure a diverse range of perspectives and expertise – where possible, recruit people with diverse skills and experience, demographics, and lived and/or living experience (without making disclosure a requirement)[57]
- enable team members to support each other and build on each other's ideas
- reduce tokenism.

It can be helpful to have people with lived and/or living experience **involved in the assessment process** as they will often have a good understanding of the context of candidates' previous experience and the expertise they will bring to the evaluation. This can also help to further integrate existing lived and/or living experience team members.

INDUCTION, ONBOARDING AND RELATIONSHIP BUILDING

To support the induction process, an **information pack** can be helpful to clearly outline:

- the team, project, and role and responsibilities
- training and support provided
- time commitment and remuneration
- key contacts in the team.

In an **induction meeting**, you can then talk through this document with the team, answer any of their questions, and spend some time getting to know each other. **Relationship building** throughout the project helps people feel part of the team, and comfortable asking questions, leaning on each other for support, and speaking up if things become overwhelming and they need to take a break or if things are confusing or unclear.

During onboarding, it is also important to determine people's communication preferences, availability, areas of interest in the evaluation, and any adjustments they feel they may need.

TRAINING AND SKILLS BUILDING

Depending on the training and experience of the lived and/or living experience team, additional **training** in evaluation methods (e.g., interviewing/ facilitation skills, survey design and analysis) might be needed. It is also important to ask lived and/or living experience team

[57] Sinclair A, Gillieatt S, Fernandes C & Mahboub L. (2023). Inclusion as Assimilation, Integration, or Co-optation? A Post-Structural Analysis of Inclusion as Produced Through Mental Health Research on Peer Support. Qualitative Health Research 33(6):543-555. doi:10.1177/10497323231163735

members about the feedback they would like on their input to develop their skills – as individuals will have different preferences. Ongoing investment in people with lived and/or living experience helps to build their skills and develops and supports ongoing relationships, which can be beneficial for both people with lived and/or living experience and the evaluation team.

It is essential that all evaluation team members undertake trauma-informed training. While some may have done this in the past, it can be particularly useful to have the whole team undertake trauma-informed training tailored to the evaluation, as a way of understanding the risks specific to the evaluation and collectively committing to an approach.

Training for lived and/or living experience researchers

People with lived and/or living experience participating in roles on an evaluation team will bring their own skills and expertise from previous roles they have been in. For example, they may have past experience in a learning and development role, which will assist with facilitating consultations and workshops, or they may be researchers in other fields or have been involved on advisory committees.

This, coupled with on-the-job training (see above), should be sufficient to equip people with lived and/or living experience to participate as members of an evaluation team. However, there are also some courses within the sector that may help to prepare people for these roles, including:

- Roses in the Ocean's range of capacity building workshops and resources
- VMIAC's Lived and Living Experience Workforce Program
- The Mental Health Coalition of SA (MHCSA) Lived Experience Workforce Program (LEWP), which provides training and resources for people with lived and/or living experience who wish to become peer workers, and connections to other organisations and opportunities that will help them build skills in lived and/or living experience advocacy (e.g., to influence systemic change), and lived and/or living experience representation (e.g., sitting on committees and panels)
- The ALIVE National Centre for Mental Health Research Translation's Lived-Experience Collective Training (e.g., the One Day Short Course – Orientation to Applied Research for Lived-Experience Researchers):
- Lived Experience Australia's range of courses to equip people with lived and/or living experience for a variety of roles (and organisations to support them in these roles):

A number of organisations also plan to/ are in the process of developing training that would support people with lived and/or living experience to participate in evaluation roles.

ONGOING EMOTIONAL AND PRACTICAL SUPPORT

Evaluations in the mental health and suicide prevention sectors often deal with difficult content, which may bring up things that people didn't expect, or lead to vicarious trauma – the distress, dissatisfaction, hopelessness, and serious mental and physical health problems that can arise from ongoing exposure to other people's trauma. It is important to put in place emotional and practical supports for **all team members** to access, if needed.

It is also important to be clear from the beginning that people can change their involvement or withdraw at any time, and to pay particular attention to team members' wellbeing throughout.

Some forms of **internal support** that can be provided include the following:

- **Preparation**
  - meetings ahead of project activities, particularly data collection, to prepare for the task and their own potential reactions
  - communication outside of meetings so everyone is comfortable with and clear about project activities
  - established response protocols for direct data collection with stakeholders.

- **Debriefing**
  - team debriefing: to reflect as a team on project activities, share learnings, and continuously improve data collection processes
  - individual debriefing: offering team members the ability to request an individual debrief, especially during data collection if anything raises concerns for them.

- **Team check-ins**
  - regular communication to ensure lived and/or living experience members are engaged, feel included, and have adequate support and opportunity to be involved – some may desire this more than others, so it's important to understand individual preferences
  - regular meetings with the whole team during active phases of data collection to reflect on what is going well, what is proving challenging, whether the evaluation is on track, individual capacity, and identifying areas of data that team members may wish not to engage in, especially if the content is sensitive or may reveal distressing content
  - opportunities for feedback from lived and/or living experience team members on how the evaluation team can better engage with lived and/or living experience team members.

Some forms of **external support** that could be provided include the following:

- **Employee Assistance Program (EAP)** – ensure all team members are aware they can contact the organisation's EAP for debriefing and support. If possible, set this up so there is tailored support for lived and/or living experience team members.
- **External supervision or mentoring** – offer this to lived and/or living experience team members if they require or desire it during key periods of the evaluation (e.g., data collection).
- **Broader services and supports within the local area** – for example via *Head to Health*.

INVOLVEMENT IN PROJECT PHASES

There are multiple ways people with lived and/or living experience can be involved in **all phases of the evaluation** to ensure the approach is sensitive and ethical for people with lived and/or living experience, the evaluation assesses what is important to people with lived

and/or living experience, and findings are interpreted from a lived and/or living experience perspective.

It is important to give lived and/or living experience team members **choice in what tasks they would like to be involved in** (based on their skills, interests, and time available for the evaluation), **and how they contribute to these tasks** (e.g., online meetings, written contributions).

Some ways that lived and/or living experience team members can be involved in project phases include the following:

- **Planning**
  - advising on and developing a program logic for the evaluation
  - advising on and developing the approach to data collection to ensure it is sensitive and trauma-informed for people with lived and/or living experience
  - advising on and developing data collection tools (e.g., surveys, interview guides).

- **Data collection**
  - conducting interviews or co-facilitating group interviews with people with lived and/or living experience (and other stakeholders if they desire)
  - co-facilitating workshops
  - conducting structured observations (i.e. observing program sessions to assess accessibility, facilitator competency, and participant engagement).

- **Analysis**
  - qualitative analysis – identification of themes, and coding data
  - quantitative analysis – analysing survey data.

- **Synthesis and reporting**
  - reflecting on key findings and helping the team interpret them in context
  - report writing
  - reviewing reports, or key sections of reports (e.g., service user experiences and outcomes), and providing feedback, for example, on whether key points are covered, the communication is accessible, and the language appropriate.

It is important to keep those who are not involved in particular tasks updated on how they are progressing, so they remain engaged and feel part of the team.

OFFBOARDING AT THE END OF THE EVALUATION

It is essential to offboard lived and/or living experience team members from the evaluation project to thank them for their time and contributions, reflect and learn as a team, and ensure lasting relationships. This also includes creating opportunities for feedback and clarity about how their contributions will be considered.

Some options to consider at this stage include:

- holding a meeting with the whole team to reflect on how the project went, thank people for their contribution, celebrate completion, and identify learnings
- gathering feedback from lived and/or living experience team members about the project and their involvement through an anonymous survey, or an exit interview/ conversation if they prefer
- gauging lived and/or living experience team members' interest to be involved in future opportunities
- explaining how lived and/or living experience team members can use their involvement in the project on their CVs and resumes
- clarifying whether there are expected to be more updates in the future (e.g., about how the evaluation has been used).

Resources

- Krysinska, K., Ozols, I., Ross, A. et al (2023). Active involvement of people with lived experience of suicide in suicide research: a Delphi consensus study. *BMC Psychiatry* 23, 496.
- Jo Farmer (2018). Evaluating with a mental health lived experience.
- Suomi, A., Freeman, B., & Banfield, M. (2018). Framework for the engagement of people with a lived experience in program implementation and research: Review and report prepared for the LifeSpan suicide prevention project. The Australian National University, Black Dog Institute.
- Hodges, E., Leditschke, A., Solonsch, L. (2023). The Lived Experience Governance Framework: Centring People, Identity and Human Rights for the Benefit of All. Prepared by LELAN (SA Lived Experience Leadership & Advocacy Network) for the National Mental Health Consumer and Carer Forum and the National PHN Mental Health Lived Experience Engagement Network. Mental Health Australia, Canberra.
- Webb, M., Cooper, C., Hemming, L., Dalton, A., Unity, E., Simmons, M., Bendall, S. & Robinson, J. (2023). Guidelines for involving young people with lived and living experience of suicide in suicide research. Melbourne: Orygen.
- Roses in the Ocean (n.d.). Resources for organisations.
- Roses in the Ocean (2022). Lived Experience of Suicide Engagement, Partnership and Integration (LESEPI) Implementation Toolkit.
- Mental Health Commission of New South Wales. (2018). Lived Experience Framework.
- Department of Health and Human Services Victoria. (2019). Mental health lived experience framework.
- Roper, C., Grey, F., and Cadogen, E. (2018). Co-production: Putting principles into practice in mental health contexts.

APPENDIX 4 EVALUABILITY ASSESSMENT CHECKLIST[58]

Focus	Question	Response
Design		
Related to need	Is the project objective clearly relevant to the needs of the target group, as identified by any form of situation analysis, baseline study, or other evidence and argument?	
Clear logic	Are the ultimate outcomes clearly identified, and are the proposed steps towards achieving these clearly defined? (logic model in place)	
Plausible logic	Is it likely that the outcomes could be achieved, given the planned programs, within the available time? Is there evidence from elsewhere that it could be achieved?	
SMART indicators	Are there specific, measurable, achievable, relevant, and time-bound indicators aligned with the logic?	
Contextual factors identified	Have assumptions about the roles of factors besides the program been made explicit? Are these factors able to be monitored?	
Information availability		
Documentation complete	Is there sufficient documentation of design and implementation and changes made, e.g., Project proposal, Progress reports?	
Baseline data available	If baseline data is not yet available, are there specific plans for when baseline data would be collected, and how feasible are these? If baseline data exists, is the raw data available? Is the sampling process clear and data collection tools/ definitions clear?	
Range of data available	To what extent does available data cover the range of outputs and outcomes?	

[58] Substantially revised drawing from Davies, R. (2013). Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations. Report of a Study Commissioned by the Department for International Development, p19-23. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf; http://www.unodc.org/documents/evaluation/Guidelines/Evaluability_Assessment_Template.pdf

Focus	Question	Response
Completeness and quality of data	Is there missing data? Are there issues with consistency?	
Clarity about stakeholders	Is there a record of who was involved in what project activities, and when?	
Practicality and utility		
Stakeholders can be engaged	Are contact details available? Was consent to contact for monitoring, research, and evaluation gained (or does this need to be sought in this project)? Are staff and key stakeholders likely to be present, or absent on leave or secondment? Can reported availability be relied upon?	
Sufficient resources	Is the funding or staff capability sufficient for the required evaluation activities? If no, can the evaluation be re-scoped (e.g., reduce the number of questions or the breadth of data collection) to make this feasible?	
Appropriate timing	Is there an opportunity for an evaluation to have an influence in the current context (e.g., competing priorities)? If the evaluation was planned in advance, is the evaluation still relevant? Are the questions relevant to priorities?	
Clear intended users	Are the intended primary users clear? Are there any secondary intended users? Can they be involved in ensuring the evaluation will meet their needs?	
Stakeholder interests	Are there clear questions for the evaluation of interest to stakeholders? If no, can these be defined with intended users (see above?) Are questions realistic, given the project design and likely data availability? If no, can they be re-scoped?	
Ethics and risk are manageable	Have ethical considerations been identified? Are they manageable? Have risks been identified? Are they manageable through identified strategies?	

APPENDIX 5 OUTCOMES MAPPING

A broad overview of the types of measures and the associated outcomes collected through the National Outcomes and Casemix Collection (NOCC), and the Primary Mental Health Care Minimum Data Set (PMHC-MDS), is outlined below.

Source NOCC	PMHC - MDS	Measure	Age group: Children & Adolescent	Age group: Adult	Administered by	Intended use	Notes
X		Health of the Nation Outcome Scale (HoNOS/ HoNOS65+)		X	Clinician, could be self-reported. Free to use	12 x 1 item scales measuring behaviour, impairment, symptoms, and social functioning for those in the 18-64 years old age group, or 65+	Potentially complex instructions as behaviours can overlap between the 12 items, different wording for each item, but all follow 0-4 item, with 9=unknown. Scoring 0-48, although interpretation on individual items for change in that domain against differing timepoints. Different time periods based on inpatient or not, so not comparable and/or potential for error. More global information than mental health and psychological distress, and includes some SDOMH.
X		Health of the Nation Outcome Scale for Children and Adolescents (HoNOSCA)	X		Clinician, can be clinician/ parent 5-18 years old, or self if 13-18 years old. Free to use	Up to 15 1 item scales measuring behaviour, impairment, symptoms, and social functioning for uchildren	Period rated not defined for most items. Same scale as HONOS.

Source NOCC	PMHC - MDS	Measure	Age group: Children & Adolescent	Age group: Adult	Administered by	Intended use	Notes
X		Abbreviated Life Skills Profile (LSP-16)		X	Clinician, can be self-reported	16 items measure of function and disability in people with mental illness (16 item scale covers Self-care, Anti-social, Withdrawal, Compliance) 18-64 years	Designed to be conversational. 4-point scale + 'unable to rate' and 'unknown.' Lower scores = higher functioning. Can be used in conjunction with HONOS. METeOR missing question 5, possibility for confusion.
X	X	Kessler Psychological Distress Scale (K10)		X	Self-report or clinician. Free to use (supposed to acknowledge author)	Questions level of nervousness, agitation, psychological fatigue, and depression during last 4 weeks Scoring 1-4, higher scores = higher impairment	Used in national mental health survey. Easy to administer, can be used for variety of purposes for global psychological distress. • score under 20 are likely to be well • score 20-24 are likely to have a mild mental disorder • score 25-29 are likely to have moderate mental disorder • score 30 and over are likely to have a severe mental disorder.

Source NOCC	PMHC - MDS	Measure	Age group: Children & Adolescent	Age group: Adult	Administered by	Intended use	Notes
	X	Kessler-5 (K5; for Aboriginal and Torres Strait Islander people if considered more appropriate)		X	Self-report or clinician. Free to use (supposed to acknowledge author)	Last 4 weeks Scoring 1-4, higher scores = higher impairment, range 5-25	There is also the K-6, however the question about feeling worthless is removed in K5 for appropriateness for First Nations.
X		Behaviour and Symptom Identification Scale (BASIS-32)		X	Typically, BASIS-32® is given at admission and discharge for hospital-based episodes of care, and at intake/ initiation of treatment and then periodically thereafter in ambulatory care settings Scoring	The BASIS-32® measures the change in self-reported symptom and problem difficulty over the course of treatment The questions are grouped into 5 domains, representing: • relation to self and others; • daily living and role functioning; • depression and anxiety;	The survey measures the degree of difficulty experienced by the patient during a one-week period on a five-point scale ranging from no difficulty to extreme difficulty. BASIS-32® assesses treatment outcomes from the patient perspective. Is the predecessor of the BASIS-24®, continues to be a leading behavioural health assessment tool. It is well-known and highly esteemed for its comprehensive outcomes measurement capability, as well as for its grounding in scientific methods of instrument development and validation.

Source NOCC	PMHC - MDS	Measure	Age group: Children & Adolescent	Age group: Adult	Administered by	Intended use	Notes
					0 - No difficulty 1 - A little difficulty 2 - Moderate difficulty 3 - Quite a bit of difficulty 4 - Extreme difficulty Scoring of this measure is conducted through a proprietary, licenced algorithm	• impulsive and addictive behaviour; • psychosis	
X		Mental Health Inventory (MHI-38)		X	Consumer self-report tool	The MHI-38 measures general psychological distress and wellbeing and includes positive aspects of wellbeing (such as cheerfulness, interest in and enjoyment of life) as	Questions are asked of the consumer about the past month, with a single allowable response for each question. The response scales are from 1-6, noting different labels are used for different questions on the form (such as

Source NOCC	PMHC - MDS	Measure	Age group: Children & Adolescent	Age group: Adult	Administered by	Intended use	Notes
						well as negative aspects of mental health (such as anxiety and depression)	"Always", "All of the time", "Extremely so", and "Never", "None of the time", "Not at all"). The two exceptions to this are questions 9 and 28, which measure from 1-5 only. Lower numbers denote that the consumer response to the question is more positive, with higher numbers denoting a more negative response. For some of the questions, a score of 1 will denote better mental health, and for some questions a score of 1 will denote poorer mental health.
X	X	Strengths and Difficulties Questionnaire (SDQ)	X		Self	'internalising problems' (emotional+peer symptoms, 10 items) 'externalising problems' (conduct+hyperactivity	Self-report focused on behaviour change and appearance of difficulties, not connected to social determinants or distress factors.

Source NOCC	PMHC - MDS	Measure	Age group: Children & Adolescent	Age group: Adult	Administered by	Intended use	Notes
						symptoms, 10 items) 'prosocial scale' (5 items)	
X		Living in the Community Questionnaire (LCQ)		X	Self, with assistance of support worker	Measurement of social inclusion for consumers of public-funded mental health services	Focused on recovery lens, as well as ways in which individuals receive care in the community. Noted as a complicated screening tool
X		Factors Influencing Health Status (FIHS; not an outcomes measure but important to interpret outcomes data)		X	Consumers with clinician input		Presence of one or more factors impacting on the relationship between social interaction/ environment with behaviour and thoughts that have a negative effect on an individual's psychological health and requires additional clinical input. Explores Social Determinants of Health and lifespan impacts.

APPENDIX 6 EMMIE TEMPLATE[59]

A modified version of the EMMIE framework may provide a useful starting point for sharing information to continue building an evidence base of what works in varying contexts. This framework was developed for policymakers seeking practical evidence for decision-making about social policy and crime prevention programs to provide a consistent means of reviewing and extracting information from different types of evaluations.

TABLE 13. TEMPLATE FOR SHARING INFORMATION ABOUT EVALUATIONS

Element	Description	Response
E the effectiveness or outcomes of an evaluation	Where relevant this could include the effect direction and size of an outcome and the confidence that should be placed on that estimate	• *Effect size* • *Exploration of the relationship between two variables* • *Measurement or consideration of unanticipated effects*
M the means by which the outcomes were achieved	The way in which change occurred when implementing the policy, practice or program in question	• *Map of possible ways change occurs - how the policy/ practice/ program facilitates change – (developed before or after analysis)* • *Assessment or statements of most likely mechanisms, and any contextual conditions (these can be narratives)*
M the moderators or contexts relevant to the evaluation	What contributed to the outcomes to vary for different people in different places	• *Analysis testing the differences that context makes to outcome (theoretically driven)* • *Analysis of data* • *Statements qualifying contextual variations (these can be narratives)*
I the insights into the success and failure in implementation	Consider the policy, practice or program, and system-level factors	• *Statements such as key components necessary for implementation* • *Statement of key components necessary for replication* • *Enablers and barriers to implementation* • *Lessons for future programs*

[59] Johnson, S., Tilley, N. & Bowers, K. (2015). Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews. Journal of Experimental Criminology. 11 459–473. https://doi.org/10.1007/s11292-015-9238-7

Element	Description	Response
E the economic costs (and benefits) associated with the policy, practice or program	Varies depending on approach	• *Quantification of program inputs* • *Quantification of program outputs* • *Quantification of intensity (e.g., spend per person)* • *Estimate of cost of implementation (and by sub-group)* • *Estimate of cost-effectiveness per unit output and by sub-group)* • *Estimate of cost-benefit (and by sub-group)*