

Evaluation of Australian Government Pilot Programs

Department of Health and Aged Care

Department of Veterans' Affairs

Department of Home Affairs

© Commonwealth of Australia 2024

ISSN 1036–7632 (Print)

ISSN 2203–0352 (Online)

ISBN 978-1-76033-935-7 (Print)

ISBN 978-1-76033-936-4 (Online)

Except for the content in this document supplied by third parties, the Australian National Audit Office logo, the Commonwealth Coat of Arms, and any material protected by a trade mark, this document is licensed by the Australian National Audit Office for use under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 Australia licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/au/>.

You are free to copy and communicate the document in its current form for non-commercial purposes, as long as you attribute the document to the Australian National Audit Office and abide by the other licence terms. You may not alter or adapt the work in any way.

Permission to use material for which the copyright is owned by a third party must be sought from the relevant copyright owner. As far as practicable, such material will be clearly labelled.

For terms of use of the Commonwealth Coat of Arms, visit the *Australian honours system* website at <https://www.pmc.gov.au/honours-and-symbols/australian-honours-system>.

Requests and inquiries concerning reproduction and rights should be addressed to:

Chief Operating Officer
Corporate Management Group
Australian National Audit Office
GPO Box 707
Canberra ACT 2601

Or via email:

communication@anao.gov.au.



Canberra ACT
17 June 2024

Dear President
Dear Mr Speaker

In accordance with the authority contained in the *Auditor-General Act 1997*, I have undertaken an independent performance audit in the Department of Health and Aged Care, Department of Veterans' Affairs and Department of Home Affairs. The report is titled *Evaluation of Australian Government Pilot Programs*. Pursuant to Senate Standing Order 166 relating to the presentation of documents when the Senate is not sitting, I present the report of this audit to the Parliament.

Following its presentation and receipt, the report will be placed on the Australian National Audit Office's website — <http://www.anao.gov.au>.

Yours sincerely



Rona Mellor PSM
Acting Auditor-General

The Honourable the President of the Senate
The Honourable the Speaker of the House of Representatives
Parliament House
Canberra ACT

AUDITING FOR AUSTRALIA

The Auditor-General is head of the Australian National Audit Office (ANAO). The ANAO assists the Auditor-General to carry out their duties under the *Auditor-General Act 1997* to undertake performance audits, financial statement audits and assurance reviews of Commonwealth public sector bodies and to provide independent reports and advice for the Parliament, the Australian Government and the community. The aim is to improve Commonwealth public sector administration and accountability.

For further information contact:
Australian National Audit Office
GPO Box 707
Canberra ACT 2601

Phone: (02) 6203 7300
Email: ag1@anao.gov.au

Auditor-General reports and information about the ANAO are available on our website:
<http://www.anao.gov.au>

Audit team

Sarina Lacey
Maurice Gauder
Dennis McCarthy
Tomas Berry
Jennifer Canfield
Barbara Das
Alexandra Collins

Contents

Summary and recommendations.....	7
Background	7
Conclusion	8
Supporting findings.....	9
Recommendations.....	11
Summary of entity responses	12
Key messages from this audit for all Australian Government entities	14
Audit findings.....	15
1. Background	16
Introduction	16
Rationale for undertaking the audit	17
Selected Australian Government pilots	17
Audit approach	22
2. Governance arrangements.....	24
Do the entities have a fit-for-purpose approach to choosing when and what to evaluate?	25
Do the entities have an appropriate framework for conducting program evaluations?	35
3. Evaluation approach.....	38
Was there appropriate evaluation planning before beginning the pilot?	39
Are evaluation criteria methodologies consistent with good practice?.....	47
4. Reporting and advice to the Australian Government	54
Was there fit-for-purpose analysis and reporting of pilot evaluation outcomes?	55
Did advice to government and policy proposals draw on pilot evaluation outcomes?	67
Appendices	71
Appendix 1 Entity responses	72
Appendix 2 Improvements observed by the ANAO	76
Appendix 3 Governance arrangements to support evaluation of pilots	77
Appendix 4 Commonwealth Evaluation Principles	78
Appendix 5 Evaluation approach and reporting and advice to the Australian Government	79



Audit snapshot

Auditor-General Report No.39 2023–24

Evaluation of Australian Government Pilot Programs



Why did we do this audit?

- ▶ Effective evaluation capability, capacity, and culture are critical in ensuring the proper use of public money.
- ▶ This audit selected five pilot programs across three Australian Government departments to review evaluation systems and processes, including how the evaluation was used to decide whether a proposed policy should be adopted.



What did we find?

- ▶ Two of the three departments have largely effective governance arrangements to support the evaluation of pilot programs.
- ▶ Largely appropriate planning and largely robust evaluation approaches were used in two out of the five pilots.
- ▶ The evaluation reports for three out of five pilots were largely comprehensive. Advice provided to the Australian Government partly reflected the learnings from the evaluation.



Key facts

- ▶ The Department of Health and Aged Care was allocated \$10 million for the Naxolone pilot and \$1.9 million for the Kava pilot.
- ▶ The Department of Veterans' Affairs was allocated \$4 million for the Wellbeing and Support Program pilot, and \$2.3 million for the Non-Liability Rehabilitation pilot.
- ▶ The Department of Home Affairs developed and implemented the Skilled Refugee Labour Agreement pilot within its existing funding allocation.



What did we recommend?

- ▶ One recommendation for the Department of the Treasury aimed at strengthening evaluation guidance.
- ▶ Five recommendations for one or more of the departments aimed at strengthening evaluation policies and improving advice to the Australian Government on the outcomes of evaluations.
- ▶ The entities agreed to all recommendations.

2–3 years

was the duration of the pilots examined during the audit.

4 out of 5

pilots examined had evaluation methodologies partly or largely consistent with the Commonwealth Evaluation Toolkit.

Summary and recommendations

Background

1. Evaluation is a structured assessment of the value of government programs or activities, aimed at supporting improvement, accountability, and decision-making throughout the policy cycle. Pilot programs are small-scale tests or trials of programs with the aim of informing future decision-making.
2. The *Public Governance, Performance and Accountability Act 2013* (the PGPA Act) requires the accountable authority of a Commonwealth entity to measure and assess the performance of the entity in achieving its purposes¹, and that a minister must not approve expenditure unless satisfied that the expenditure would be an efficient, effective, economical and ethical use of public money.²
3. In 2019, the Australian Government released the *Independent Review of the Australian Public Service*³, which recommended that the APS embed a culture of evaluation and learning from experience to underpin evidence-based policy and delivery (Recommendation 26). The Australian Government agreed in part to this recommendation.⁴ The Minister for Finance endorsed a Commonwealth Evaluation Policy⁵ and Resource Management Guide 130 Commonwealth Evaluation Toolkit⁶ (the Toolkit) on 1 December 2021. The Toolkit provides a principles-based approach for the conduct of evaluations. It applies to all Commonwealth entities and companies subject to the PGPA Act.

Rationale for undertaking the audit

4. Pilot programs are trial programs of limited size that are used to decide whether a proposed policy should be adopted, and what adjustments should be made before adoption. Monitoring and evaluation are critical components of a pilot to support an assessment of the program or activity's impact and efficiency.
5. The audit involved the examination of five Australian Government pilot programs across the Department of Health and Aged Care (Health), the Department of Home Affairs (Home Affairs), and the Department of Veterans' Affairs (DVA). The pilots ranged in length from two to

1 *Public Governance, Performance and Accountability (PGPA) Act 2013*, section 38.

2 *Public Governance, Performance and Accountability (PGPA) Act 2013*, subsection 71(1).

3 Department of the Prime Minister and Cabinet, *Independent Review of the Australian Public Service*, PM&C, Canberra, 2018, available from <https://www.pmc.gov.au/sites/default/files/resource/download/independent-review-aps.pdf> [accessed 15 January 2024].

4 The Australian Government did not agree to the recommended systematic changes to Cabinet and Budget advice processes. See: The Senate: *APS Inc: undermining public sector capability and performance*, 2021 Appendix 1: Government response to the recommendations of the Independent Review of the APS, available from [https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/CCurrentAPSCapabilities/Report](https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/CurrentAPSCapabilities/Report) [accessed 2 April 2024].

5 Australian Centre for Evaluation, *Commonwealth Evaluation Policy*, available from <https://evaluation.treasury.gov.au/about/commonwealth-evaluation-policy> [accessed 1 April 2024].

6 Australian Centre for Evaluation, *Commonwealth Evaluation Toolkit*, available from <https://evaluation.treasury.gov.au/toolkit/commonwealth-evaluation-toolkit> [accessed 1 April 2024].

three years. The audit provides assurance to the Parliament over the appropriateness of frameworks for evaluation, and the adequacy of evaluation of pilot programs.

Audit objective and criteria

6. The objective of the audit was to assess the effectiveness of the evaluation of selected Australian Government pilot programs.
7. To form a conclusion against this objective, the following high-level criteria were adopted:
 - Do the selected entities have governance arrangements in place to support effective program evaluation?
 - Was the evaluation approach for the selected pilot programs robust?
 - Was pilot program reporting and advice to government appropriate?

Conclusion

8. The evaluation of the selected Australian Government pilot programs was mixed. Health's evaluation of the Take Home Naloxone pilot was largely effective, and the evaluation of the Kava pilot was partly effective. DVA's evaluation of the Wellbeing and Support Program pilot was largely effective, and the evaluation of the Non-Liability Rehabilitation pilot was partly effective. Home Affairs' evaluation of the Skilled Refugee Labour Agreement pilot was partly effective.

9. Health and DVA have largely effective governance arrangements to support the evaluation of pilot programs. Home Affairs has partly effective arrangements. Health and DVA have strengthened their governance arrangements through the updating or development of entity-specific frameworks, guidance, and training on what, when and how to conduct an evaluation. Home Affairs does not have entity-specific evaluation guidance. Evaluation culture is maturing within Health and DVA, and is immature at Home Affairs. Pilot programs are only referenced in DVA's entity-specific guidance.

10. The evaluation planning and approach for Health's Take Home Naloxone pilot and DVA's Wellbeing and Support Program pilot were largely robust, including appropriate stakeholder engagement and relevant ethics approvals. Planning for the evaluation of Health's Kava pilot did not identify the risk that ethics approval may not be granted for one of the planned qualitative analysis methods, and there was a lack of baseline evidence to support the planned evaluation methodology. The effectiveness of planning for the evaluation of DVA's Non-Liability Rehabilitation pilot was reduced as the analytical methodologies were not documented, and no external stakeholders were consulted. Home Affairs did not complete its planning for, or undertake, a robust evaluation for the Skilled Refugee Labour Agreement pilot. All evaluation plans and approaches could have been enhanced by a greater focus on the availability of data and an assessment of the proper use of public money.

11. Health's evaluation reporting and advice to the Australian Government for the Take Home Naloxone pilot was largely effective, with the recommendations made to expand the naloxone pilot largely informed by the lessons learnt from the evaluation. Health's evaluation reporting and advice to the Australian Government for the Kava pilot was partly effective as neither the evaluation report nor recommendations on the continuation of the pilot have been provided to the Australian Government. The evaluation report for DVA's Wellbeing and Support Program pilot

was largely effective. There was no evidence of DVA advising the Australian Government on the evaluation findings and impact on future program design. The evaluation for the Non-Liability Rehabilitation pilot has not yet commenced, and reporting and advice to the Australian Government on the mid-pilot review was partly effective. Home Affairs' evaluation reporting and advice to the Australian Government for the Skilled Refugee Labour Agreement pilot was partly effective, with outputs rather than pilot outcomes analysed and reported to the minister.

Supporting findings

Governance arrangements

12. The Commonwealth Evaluation Toolkit provides appropriate high-level guidance to support entities in determining what programs or policies should be evaluated and when. It provides limited guidance on conducting an economic evaluation, including any assessment of cost effectiveness of implementation, and does not include a requirement for all pilots to be evaluated.

- With the exception of DVA's Non-Liability Rehabilitation pilot, at the time the other examined pilots commenced, only Health had established internal evaluation guidance.
- In November 2023, Health published a revised evaluation strategy which specifies roles and responsibilities and includes a tiered system for identifying evaluation priorities across the department.
- Since the commencement of the Wellbeing and Support Program pilot, DVA developed a framework which supports when and what to evaluate based on program characteristics, timing and capability. In August 2023, DVA introduced a framework for the planning, monitoring and evaluation of its health and wellbeing programs, which includes roles and responsibilities.
- Home Affairs does not have an entity-specific approach to determining when and what to evaluate.
- Each entity has an internal evaluation team to provide guidance and support on evaluation practice.

(See paragraphs 2.7 to 2.58)

13. Health and DVA have policies and guidance materials for how to conduct program evaluations. Only Health has guidance on when economic evaluation should be undertaken and the guidance is limited. Training on evaluation practices is provided at Health and DVA. Attendance is not consistently monitored. Home Affairs has no entity-specific guidance on conducting evaluations, and no training programs available to staff. (See paragraphs 2.59 to 2.69)

Evaluation approach

14. Planning for evaluation, including stakeholder engagement, was completed for Health's Take Home Naloxone pilot and DVA's Wellbeing and Support Program pilot. Planning for stakeholder engagement for evaluation of Health's Kava pilot did not account for the risk that ethics approval may not be granted and the resulting impact on the planned analysis and evaluation methodology. The effectiveness of planning for the evaluation of DVA's Non-Liability Rehabilitation pilot was reduced as the methodologies to be used were not documented, and no

external stakeholders were consulted. Home Affairs did not complete its planning for the evaluation for the Skilled Refugee Labour Agreement pilot. While data sources were identified within the evaluation plans that were developed, one or more planned data sources within each pilot were not available for the evaluation, and this risk had not been identified. (See paragraphs 3.6 to 3.52)

15. The evaluation methodologies used for three out of the five pilots examined were largely consistent with the Toolkit. For the evaluations conducted, all could have been strengthened with a greater focus on baseline data, control group outcomes, and an assessment of the proper use of public money. Ethics approvals were obtained for Health's Take Home Naloxone pilot and DVA's Wellbeing and Support Program pilot. The ethics approval sought for Health's Kava pilot was not granted and no alternative strategy was developed to obtain information that was critical to the evaluation. DVA's Non-Liability Rehabilitation pilot evaluation plan did not include a consideration of ethics approval and the post-implementation review has not yet been undertaken. As Home Affairs did not conduct an evaluation of its pilot, there was no methodology applied, or consideration of the need for ethics approval. (See paragraphs 3.53 to 3.80)

Reporting and advice to the Australian Government

16. The analysis of pilot evaluation outcomes for Health's pilots and DVA's Wellbeing and Support Program pilot were largely fit for purpose, with the evaluation reports documenting the application of statistical methods to provide defensible findings and make recommendations on the basis of the analysis completed. The evaluation of DVA's Non-Liability Rehabilitation pilot has not yet commenced. Home Affairs' reporting of outputs of the Skilled Refugee Labour Agreement pilot did not contain fit-for-purpose analysis and does not satisfy the requirements of evaluation reporting in the Commonwealth Evaluation Toolkit. (See paragraphs 4.5 to 4.33)

17. Advice provided by Health to the Australian Government in relation to the Take Home Naloxone pilot was appropriate, including the lessons learnt from the pilot. The recommendation to expand the pilot into different environments was partly informed by evaluation. Health has not provided advice to the Australian Government on the findings of the evaluation or lessons learnt in relation to the Kava pilot. DVA did not advise the Minister for Veterans' Affairs on the evaluation findings or lessons learnt for future program delivery for the Wellbeing and Support Program pilot. Home Affairs' advice to the Australian Government for the continuation of the Skilled Refugee Labour Agreement pilot was not informed by an evaluation. (See paragraphs 4.34 to 4.52)

Recommendations

Recommendation no. 1 The Department of the Treasury update the Commonwealth Evaluation Policy and Toolkit to include:
Paragraph 2.16

- (a) a definition of a 'pilot';
- (b) guidance on how to conduct an economic evaluation and other methods for considering whether spending represents an appropriate use of public money;
- (c) a recommendation that evaluations of pilot programs be undertaken;
- (d) a recommendation for evaluation planning to be conducted alongside pilot design; and
- (e) guidance on governance arrangements for cross-entity evaluations to minimise duplication and maximise coordination and learnings across entities.

Department of the Treasury's response: *Agreed.*

Recommendation no. 2 The Departments of Health and Aged Care and Veterans' Affairs include in their entity-specific evaluation policies:
Paragraph 2.32

- (a) decision-making criteria for the appropriate style of evaluation to be completed by reference to the activity's risk, objective and outcomes;
- (b) guidance on how to demonstrate whether a program represented a proper use of public money, including the cost-effectiveness of its implementation, and how to undertake an economic evaluation where appropriate; and
- (c) guidance related to evaluation of pilot programs.

Department of Health and Aged Care's response: *Agreed.*

Department of Veterans' Affairs' response: *Agreed.*

Recommendation no. 3 The Department of Home Affairs develop entity-specific policies for evaluation, including:
Paragraph 2.35

- (a) decision-making criteria as to when an evaluation is required and the appropriate style of evaluation by reference to the activity's risk, objective and outcomes;
- (b) guidance on how to demonstrate whether a program represented a proper use of public money, including the cost-effectiveness of its implementation, and how to undertake an economic evaluation where appropriate; and
- (c) guidance related to evaluation of pilot programs.

Department of Home Affairs' response: *Agreed.*

Recommendation no. 4
Paragraph 2.55

The Departments of Health and Aged Care, Veterans' Affairs and Home Affairs develop and implement explicit guidance to support early engagement with central evaluation teams to improve evaluation strategy and planning.

Department of Health and Aged Care's response: *Agreed.*

Department of Veterans' Affairs' response: *Agreed.*

Department of Home Affairs' response: *Agreed.*

Recommendation no. 5
Paragraph 3.25

The Departments of Health and Aged Care, Veterans' Affairs and Home Affairs ensure evaluation plans are prepared for policies or programs subject to evaluation requirements and that they be approved prior to the implementation of the policy or program. Consistent with the Commonwealth Evaluation Toolkit, evaluation plans should incorporate proportionate and risk-based level of information, including:

- (a) methods for measuring or capturing baseline evidence, and attributing changes to the pilot, policy or program; and
- (b) a method of economic evaluation or other means of assessing the proper use of public money.

Department of Health and Aged Care's response: *Agreed.*

Department of Veterans' Affairs' response: *Agreed.*

Department of Home Affairs' response: *Agreed.*

Recommendation no. 6
Paragraph 4.50

The Departments of Veterans' Affairs' and Home Affairs' advice to government on the cessation, continuation or scaling up of a pilot draws on evidence and learnings from the evaluation, including limitations on the robustness of the evaluation undertaken.

Department of Health and Aged Care's response: *Agreed.*

Department of Veterans' Affairs' response: *Agreed.*

Department of Home Affairs' response: *Agreed.*

Summary of entity responses

18. The proposed audit report was provided to Health, DVA, Home Affairs and the Department of the Treasury. Letters of response provided by each entity are included at Appendix 1. The summary responses provided are included below. The improvements observed by the ANAO during the course of this audit are at Appendix 2.

Department of Health and Aged Care

The Department of Health and Aged Care welcome the findings, in the report and accept the recommendation directed to the department. The department is committed to implementing the recommendations effectively and has already taken steps to address issues identified in this audit.

The ANAO found the department has largely effective governance arrangements to support evaluation. The audit also found the department's evaluation culture is maturing, including:

- updating our guidance and training on what, when and how to conduct an evaluation.
- establishing the role of Chief Evaluation Officer to provide strategic oversight of evaluation activities and to engage with other Senior Executive to champion evaluation as part of policy design and program management.

The department notes the finding on the need to develop better guidance on conducting economic evaluations or other means of assessing the proper use of public money.

Since the audit was conducted, the department has launched its Strategic Investment Framework, which makes sure our policy and program officers embed evaluation and evidence within all programs. The Framework will ensure investments are supported by robust, evidence-based program evaluation and target funding to high-value programs aligned with priority areas.

The department notes that the audit on the Kava Pilot Program was undertaken while the pilot period was still under way, and certain aspects of the pilot, including recommendations to Government on the future of the Program, are yet to be finalised.

The department is building its in-house evaluation capability through a range of initiatives including:

- implementing the new *Evaluation Strategy 2023-26*
- developing a suite of departmental-specific tools and resources to support high-quality evaluation.
- partnering with Australian Centre for Evaluation in Treasury and leveraging opportunities to showcase in-house impact evaluation capability.

Department of Veterans' Affairs

The Department of Veterans' Affairs (DVA) welcome the ANAO recommendations. The ANAO report acknowledges that DVA has established policies and processes that largely support compliance with the Commonwealth Evaluation Policy (the Policy).

The Department acknowledge and agree with the ANAO's recommendations. Work is planned for 2024 to review and update the relevant policies and protocols to enhance maturity with the Commonwealth Evaluation Policy requirements, and work has already commenced to implement these enhancements.

Department of Home Affairs

The department agree with the recommendations, and as part of its ongoing efforts to strengthen evaluation, acknowledge the benefits of a more robust evaluation culture to inform Government decision-making.

The department continues to leverage Commonwealth resources and materials to assist in guiding staff on how an evaluation should be carried out. To supplement the Commonwealth Evaluation Toolkit, the department is developing additional resources to assist staff in determining when, and to what extent, an evaluation should be conducted.

The department is monitoring the outcomes of the Skilled Refugee Labour Agreement to build a sufficient evidence base to assess the viability and future scalability of the program. The

department's advice to Government on the future of the Skilled Refugee Labour Agreement Pilot will be informed by an evaluation consistent with the Commonwealth Evaluation Policy.

Department of the Treasury

Treasury welcomes the report and agrees with the recommendation to update guidance in the Commonwealth Evaluation Toolkit (the Toolkit). Specifically, Treasury will update the Toolkit to include a definition of a 'pilot', and provide guidance on: economic evaluation, evaluation of pilots, and governance arrangements for cross-entity evaluations.

Treasury's guidance on whether spending represents an appropriate use of public money will focus on (and be limited to) guidance on economic evaluation methods, and other fit-for-purpose evaluation approaches. The broader importance of appropriately using public money is well addressed through the suite of guidance administered by the Department of Finance to support resource management and therefore will not be duplicated through Treasury materials.

Treasury will recommend, but not mandate, that all pilots are subject to evaluation consistent with the principles-based Commonwealth Evaluation Policy, which recommends that responsible managers need to determine robust, proportional evaluation approaches for specific pilots or programs.

The Department of the Treasury is committed to continuous improvement of the Evaluation Toolkit. Planned enhancements will include more practical guidance on analytical methods, including economic evaluation, and effective governance arrangements that can help to improve the way Commonwealth entities assess implementation, measure the impact of government programs, and frame policy decisions.

Key messages from this audit for all Australian Government entities

19. Below is a summary of key messages, including instances of good practice, which have been identified in this audit and may be relevant for the operations of other Australian Government entities.

Performance and impact measurement

- Pilots provide an opportunity to effectively evaluate new or amended policy and program design and activities to ensure expenditure is an efficient, effective, economical, and ethical use of public money.
- Strong evaluation culture is needed to build effective evaluation capacity. This includes senior leaders prioritising evaluation activities, accessible and tailored guidance and tools for staff, transparently sharing lessons learnt, and acting on evidence-based outcomes and recommendations.
- Early engagement with evaluation expertise is needed to determine the appropriate type of evaluation for the policy or program, to identify and manage evaluation risks, and ensure the collection, and appropriate assessment, of data and information to draw robust policy conclusions.
- Early planning to identify and capture baseline and relevant data and information will help to support robust data analysis during the evaluation.

Audit findings

1. Background

Introduction

Program evaluation

1.1 Resource Management Guide 130 *Evaluation in the Commonwealth* (the Toolkit) states that evaluation is a structured assessment of the value of government programs or activities, aimed at supporting improvement, accountability, and decision-making throughout the policy cycle.⁷

1.2 The *Public Governance, Performance and Accountability Act 2013* (the PGPA Act) requires the accountable authority of a Commonwealth entity to measure and assess the performance of the entity in achieving its purposes⁸, and that a minister must not approve expenditure unless satisfied that the expenditure would be an efficient, effective, economical and ethical use of public money.⁹

1.3 The Budget Process Operational Rules (Rule 1.12) state:

Entities should provide sufficient detail to demonstrate that an evaluation plan will be established in the early stages of policy implementation, taking into account the Commonwealth Evaluation Principles (fit-for-purpose, useful, robust, ethical and culturally appropriate, credible and transparent where appropriate) to allow the effectiveness and efficiency of programs to be measured and assessed over time, consistent with the Commonwealth Evaluation Policy.

1.4 In 2019, the Australian Government released the *Independent Review of the Australian Public Service* (the APS Review), which recommended that the APS embed a culture of evaluation and learning from experience to underpin evidence-based policy and delivery (Recommendation 26). The Australian Government agreed in part to this recommendation.¹⁰ The Minister for Finance endorsed a Commonwealth Evaluation Policy and the supporting Toolkit, which was released and took effect on 1 December 2021. The Toolkit provides a principles-based approach for the conduct of evaluations. It applies to all Commonwealth entities and companies subject to the PGPA Act.

1.5 In October 2022 the Australian Government announced the APS Reform agenda, informed by the recommendations of the APS Review. Outcome eight of the reform agenda is for the APS to continuously improve its capabilities, including embedding a culture of evaluation. The 2023–24 Budget included \$10 million over four years for an Australian Centre for Evaluation within the Department of the Treasury. The government established the Australian Centre of Evaluation to:

strengthen the quantity, quality, and impact of evaluations across the APS and to lead the APS to integrate high-quality evaluation into all aspects of program and policy development to support

7 Australian Centre for Evaluation, *Commonwealth Evaluation Toolkit*, available from <https://evaluation.treasury.gov.au/toolkit/when-evaluate> [accessed 1 April 2024].

8 *Public Governance, Performance and Accountability (PGPA) Act 2013* section 38.

9 *Public Governance, Performance and Accountability (PGPA) Act 2013* subsection 71(1).

10 The Australian Government did not agree to the recommended systematic changes to Cabinet and Budget advice processes. See: The Senate: *APS Inc: undermining public sector capability and performance*, 2021 Appendix 1: Government response to the recommendations of the Independent Review of the APS, available from [https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/CCurrentAPSCapabilities/Report](https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/CurrentAPSCapabilities/Report) [accessed 2 April 2024].

evidence-based policy decisions that deliver better outcomes for Australians. Doing so enables agencies to better identify whether programs or policies are meeting their intended outcomes.¹¹

Pilot programs

1.6 Pilot programs or activities enable evaluation through small-scale tests or trials of aspects of program logic with the aim of informing future decision-making. Pilot programs are not mentioned in the Commonwealth Evaluation Toolkit, Cabinet Handbook or Budget Process Operational Rules. Pilot programs that are testing or refining an alternative design need:

- an evaluation methodology to be integrated into the pilot design;
- monitoring to ensure the pilot is generating useful evidence to support decision making;
- the evaluation of evidence relating to policy design, administration, impact and efficiency; and
- the presentation of analysis and findings to government to inform policy decisions.

Rationale for undertaking the audit

1.7 Pilot programs are trial programs of limited size that are used to decide whether a proposed policy should be adopted, and what adjustments should be made before adoption. Monitoring and evaluation are critical components of a pilot to support an assessment of the program or activity's impact and efficiency.

1.8 The audit involved the examination of five Australian Government pilot programs across the Department of Health and Aged Care (Health), the Department of Home Affairs (Home Affairs), and the Department of Veterans' Affairs (DVA). The pilots ranged in length from two to three years. The audit provides assurance to the Parliament over the appropriateness of frameworks for evaluation, and the adequacy of evaluation of pilot programs.

Selected Australian Government pilots

Department of Health and Aged Care

Take Home Naloxone pilot

1.9 Naloxone has been used in Australia in emergency settings for many years and is a treatment for temporarily reversing the effects of an opioid overdose. In relation to the Take Home Naloxone pilot (Naloxone pilot), Health advised the Australian Government that naloxone was underutilised by the target population and that key barriers to access included the cost of purchasing naloxone over the counter or via prescription and visiting a prescriber for Pharmaceutical Benefits Scheme (PBS) funded access.

1.10 The Australian Government approved the Naloxone pilot on 1 April 2019. The purpose of the Naloxone pilot was 'to reduce known barriers currently preventing a greater uptake of naloxone, allowing people to access naloxone over-the-counter, subsidised by the Government'.

11 Australian Public Service Reform Office, *Annual Progress Report 2023*, available from <https://www.apsreform.gov.au/sites/default/files/resource/download/APS%20Reform%20Annual%20Progress%20Report%201.pdf>, p.24 [accessed 1 April 2024].

1.11 To facilitate the pilot, a legislative instrument was approved by a delegate of the Minister for Health and Aged Care (the *National Health (Take Home Naloxone Pilot) Special Arrangement 2019*) to enable individuals in the trial states not to be charged a PBS co-payment for naloxone, and not be required to visit a medical practitioner or nurse practitioner, or dentist to get a prescription. As part of the pilot, naloxone was also available in a broader range of settings (beyond pharmacies) such as hospitals and from other authorised persons, or organisations, such as needle and syringe programs, alcohol and other drug treatment centres or correctional release programs.

1.12 Funding for the pilot was announced in the 2019–20 Budget, with \$7.2 million allocated over two years to establish the pilot in New South Wales and South Australia. In February 2019 funding was increased to \$10 million and Western Australia was added to the pilot for 2019–20 to 2021–22.

1.13 The pilot ran from 1 December 2019 to 30 June 2022. The University of Queensland was contracted in September 2019 to undertake the evaluation, which was completed in January 2022. The evaluation was based on the implementation of the pilot from commencement to 30 June 2021.

1.14 In the March 2022–23 Federal Budget, the Australian Government allocated \$19.6 million over four years to expand the pilot to all Australian states and territories. On 16 June 2022 the Minister for Health and Aged Care approved a national Take Home Naloxone Program following the announcement made in the Federal Budget in May 2022. The timeline for the Naloxone pilot is set out in Table 1.1.

Table 1.1: Timeline of the Naloxone pilot

Date	Event
February 2019	Pilot announced
September 2019	University of Queensland engaged to undertake evaluation
1 December 2019	Pilot commenced
30 June 2021	Conclusion of evidence period for the evaluation report
September 2021	Evaluation report provided to Health
June 2022	Pilot concluded
June 2022	Minister agreed to progress implementation of a national Take Home Naloxone program, following announcement of additional funding in the 2022–23 Budget

Source: ANAO analysis of the Department of Health and Aged Care records.

Kava pilot

1.15 On 16 January 2019, the Prime Minister announced a pilot to ease personal import limits for kava (the Kava pilot).¹² Kava is a member of the pepper plant family native to the Pacific Islands and has traditional uses as a ceremonial drink. It contains the psychoactive ingredient kavalactones and acts as a central nervous system depressant.¹³ The pilot was subsequently expanded to include

12 Department of the Prime Minister and Cabinet, *Joint Remarks with the Prime Minister of Vanuatu*, 2019, available from <https://pmtranscripts.pmc.gov.au/release/transcript-42058> [accessed 2 April 2024].

13 NSW Health, *Kava clinical factsheet* [Internet], 2022, available from <https://www.health.nsw.gov.au/aod/resources/Pages/kava-clinical-factsheet.aspx> [accessed 2 April 2024].

commercial importation of kava following a further announcement by the Prime Minister on 11 October 2019.¹⁴

1.16 The purpose of the Kava pilot was to increase stronger cultural and economic ties between Australia and Pacific Island countries.

1.17 The Australian Government allocated \$1.9 million in funding for the Kava pilot from 2020–21 to 2023–24.

1.18 The Kava pilot consisted of two phases.

- Phase One — increasing the amount of kava that incoming passengers (18 years or over) are allowed to bring into Australia in their accompanied baggage from two kilograms to four kilograms.
- Phase Two — allowing the commercial importation of kava into Australia, which was supported by changes to the *Customs (Prohibited Imports) Regulations 1956* and the Australia New Zealand Food Standards Code – Standard 2.6.3 – Kava.

1.19 The Kava pilot ended on 31 December 2023. The arrangements under the pilot remain in place. As of March 2024, the Australian Government has not made a decision on the future of the pilot.

1.20 The implementation of the pilot was coordinated by Health, with a range of Australian Government departments and state and territory governments. The timeline of the Kava pilot is set out in Table 1.2.

1.21 Two evaluations were undertaken of the Kava pilot.

- (a) Health engaged the National Drug and Alcohol Research Centre and Ninti One¹⁵, in collaboration with other partners, to monitor and evaluate the kava pilot. The evaluation assessed the health, social, cultural, and economic effects of the pilot on the Australian community, particularly in Aboriginal and Torres Strait Islander and Pacific Islander communities.
- (b) The Department of Foreign Affairs and Trade engaged the Pacific Horticultural and Agricultural Market Access Plus program to measure the impact of Phase Two of the Kava pilot on kava growing countries in the region.

Table 1.2: Timeline of the Kava pilot

Date	Event
February 2019	Prime Minister letter to ministers directing development of Phase One of pilot and evaluation
October 2019	Prime Minister announces Phase Two of pilot
December 2019	Phase One of pilot commenced
April 2021	National Drug and Alcohol Research Centre and Ninti One engaged to conduct the evaluation

14 Department of the Prime Minister and Cabinet, *Stepping up trade and cultural ties in the Pacific*, 2019, available from <https://pmtranscripts.pmc.gov.au/release/transcript-42460> [accessed 7 April 2024].

15 The National Drug and Alcohol Research Centre at University of NSW is a research group focused on the alcohol and other drugs sector. Ninti One is an Indigenous research organisation.

Date	Event
December 2021	Phase Two of pilot commenced
September 2023	Evaluation report provided to Health
31 December 2023	Pilot concluded (arrangements under the pilot remain in place)

Source: ANAO analysis of Health records.

Department of Veterans' Affairs

Wellbeing and Support Program pilot

1.22 The Department of Veterans' Affairs (DVA) developed the Wellbeing and Support Program (WASP) pilot as a response to recommendations of the Senate Inquiry into Suicide by Veterans and Ex-Service Personnel.¹⁶ The purpose of the pilot was to 'provide an intensive and supportive case management service for transitioning members and veterans who identified as needing additional support'.

1.23 The WASP Pilot provided additional support to a specific cohort of veterans including clinical care coordination, dedicated case managers and improved access to DVA and community-funded support services.

1.24 The Australian Government allocated \$4 million in funding to the pilot from 1 July 2018 to 30 June 2020 to provide case management services for up to 200 veterans. The 2021–22 Budget announced \$23.3 million over four years to continue WASP to support up to 200 veterans from 1 July 2021. Table 1.3 sets out the timeline for the pilot.

Table 1.3: Timeline of the WASP pilot

Date	Event
15 August 2017	Senate Inquiry into Suicide by Veterans and Ex-Service Personnel report, <i>The Constant Battle: Suicide by Veterans</i> , tabled in Parliament
24 October 2017	Government response to <i>The Constant Battle: Suicide by Veterans</i> , agreeing or agreeing in principle to all recommendations, tabled in Parliament
22 August 2018	Pilot commenced
October 2018	Flinders University engaged to conduct evaluation
September 2020	Pilot concluded
January 2021	Evaluation report provided to DVA
May 2021	Further \$23.3 million of funding announced for the WASP program for four years

Source: ANAO analysis of the Department of Veterans' Affairs records.

16 The Senate, *Foreign Affairs, Defence and Trade References Committee: The Constant Battle: Suicide by Veterans*, 2017, recommendations 1 and 15, available from https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Foreign_Affairs_Defence_and_Trade/VeteranSuicide/~/_media/Committees/fadt_ctte/VeteranSuicide/report.pdf [accessed 1 April 2024].

Non-Liability Rehabilitation pilot

1.25 The Non-Liability Rehabilitation pilot (the NLR pilot) was announced in the 2021–22 Budget and commenced on 1 January 2022. The purpose of the pilot was ‘to provide access to vocational and psychosocial rehabilitation without the requirement for any liability claim to be lodged’.

1.26 The NLR pilot was designed to respond to elements of Recommendation 6.3 of the Productivity Commission’s 2019 report, *A Better Way to Support Veterans*, relating to earlier access for veterans to rehabilitation supports.¹⁷ Prior to the NLR pilot, most veterans under the *Military Rehabilitation and Compensation Act 2004* and *Safety, Rehabilitation and Compensation (Defence-related Claims) Act 1988* must have had a liability claim accepted before funding by DVA for rehabilitation could be provided. The Australian Government allocated \$2.3 million in funding for the NLR pilot over four financial years from 2021–22. Table 1.4 sets out the timeline for the pilot.

Table 1.4: Timeline of the NLR pilot

Date	Event
1 January 2022	Pilot commenced
September 2022	DVA evaluation plan approved
January 2023	Mid-pilot review completed
31 December 2023	Pilot closed to new participants
31 March 2024	Pilot concluded
April 2024	Evaluation of pilot expected to commence ^a

Note a: DVA advised the ANAO in October 2023 that the evaluation is expected to commence in April 2024. At the time of writing, the evaluation had not commenced.

Source: ANAO analysis of DVA records.

Department of Home Affairs

Skilled Refugee Labour Agreement pilot

1.27 The 2019 *Review into Integration, Employment and Settlement Outcomes for Refugees and Humanitarian Entrants in Australia* recommended the introduction of an employer sponsored visa offering immediate employment opportunities to suitably skilled refugees.¹⁸

1.28 The Department of Home Affairs (Home Affairs) developed the Skilled Refugee Labour Agreement (SRLA pilot) to allow endorsed Australian businesses to enter into a labour agreement with the Australian Government to sponsor skilled refugees through a specific employer-sponsored visa program. The purpose of the SRLA pilot was ‘to provide skilled refugees with a pathway to living and working in Australia’. The pilot initially allowed for 50 to 100 places and the Minister for Home Affairs expanded it to 200 in March 2022. Table 1.5 sets out the timeline for the pilot.

17 Productivity Commission, *Compensation and rehabilitation for Veterans: A better way to support Veterans*, 2019, available from <https://www.pc.gov.au/inquiries/completed/veterans/report> [accessed 1 April 2024].

18 Shergold P., Benson K. and Piper M., *Investing in Refugees, Investing in Australia — The findings of a Review into Integration, Employment and Settlement Outcomes for Refugees and Humanitarian Entrants in Australia*, 2019, available from <https://www.homeaffairs.gov.au/reports-and-pubs/files/review-integration-employment-settlement-outcomes-refugees-humanitarian-entrants.pdf> [accessed 1 April 2024].

Table 1.5: Timeline of the SRLA pilot

Date	Event
November 2019	Minister for Immigration, Citizenship, Migrant Services and Multicultural Affairs agrees to the development of a SRLA pilot
July 2021	Pilot commenced
March 2022	Pilot expanded to 200 places
June 2023	Home Affairs advises the Minister for Immigration, Citizenship and Multicultural Affairs on the outcome and feedback on the pilot to date. The minister approves extension of pilot to 30 June 2025.

Source: ANAO analysis of Home Affairs records.

Audit approach

Audit objective, criteria and scope

1.29 The objective of the audit was to assess the effectiveness of the evaluation of selected Australian Government pilot programs.

1.30 To form a conclusion against this objective, the following high-level criteria were applied:

- Do the selected entities have governance arrangements in place to support effective program evaluation?
- Was the evaluation approach for the selected pilot programs robust?
- Was pilot program reporting and advice to government appropriate?

Audit methodology

1.31 The audit involved:

- analysing each responsible entity's documentation of the pilot development, evaluation, and advice to the Australian Government;
- assessing each responsible entity's evaluation frameworks, policies and culture, including alignment with Australian Government guidance and how they were applied to the respective pilots; and
- meetings with relevant staff from Health, DVA and Home Affairs.

1.32 Australian Government entities largely give the ANAO electronic access to records by consent, in a form useful for audit purposes. In April 2022 Health advised the ANAO that it would not voluntarily provide certain information requested by the ANAO due to concerns about its obligations under the *Privacy Act 1988*, secrecy provisions in Health portfolio legislation, confidentiality provisions in contracts and the *Public Interest Disclosure Act 2013*. For the purposes of this audit, the Auditor-General therefore exercised powers under section 33 of the *Auditor-General Act 1997* to enable authorised ANAO officers to attend premises and examine and take copies of documents. Health facilitated authorised officers attending the department's premises to examine and copy documents, however the requirement was extended by Health to all documents, including those that did not relate to the department's obligations under legislation. Health advised that this type of information largely was not segregated in the department's record-keeping systems and the department could not be certain, in providing documents through

electronic means, that documents containing this type of information were excluded. To provide comfort to the Secretary regarding the department's obligations under portfolio legislation, on 14 September 2023 the Auditor-General issued the Secretary with a notice to provide information and produce documents pursuant to section 32 of the *Auditor-General Act 1997*. Under this notice, Health complied by agreeing to provide the information and documents requested through electronic means.

1.33 The audit was conducted in accordance with ANAO Auditing Standards at a cost to the ANAO of approximately \$392,000.

1.34 The team members for this audit were Sarina Lacey, Maurice Gauder, Dennis McCarthy, Tomas Berry, Jennifer Canfield, Barbara Das and Alexandra Collins.

2. Governance arrangements

Areas examined

This chapter examines whether the selected entities had fit-for-purpose arrangements to support the evaluation of pilot programs, including guidance regarding when and what to evaluate, and appropriate frameworks for conducting program evaluations.

Conclusion

The Department of Health and Aged Care (Health) and the Department of Veterans' Affairs (DVA) have largely effective governance arrangements to support the evaluation of pilot programs. The Department of Home Affairs (Home Affairs) has partly effective arrangements. Health and DVA have strengthened their governance arrangements through the updating or development of entity-specific frameworks, guidance, and training on what, when and how to conduct an evaluation. Home Affairs does not have entity-specific evaluation guidance. Evaluation culture is maturing within Health and DVA, and is immature at Home Affairs. Pilot programs are only referenced in DVA's entity-specific guidance.

Areas for improvement

The ANAO made one recommendation to the Department of the Treasury to update the Commonwealth Evaluation Policy and Toolkit to address pilots, include further guidance on economic evaluations, how to select an appropriate evaluation type, and coordinating multi-entity evaluations. The ANAO made three recommendations to the departments examined in the audit aimed at developing or strengthening entity-specific evaluation guidance on what, when and how to evaluate, including assessing the appropriate use of public money.

The ANAO identified two areas for improvement relating to guidance on the external publication of evaluations and entity-specific guidance on data analysis.

2.1 The 2019 *Independent Review of the Australian Public Service* (the APS Review) identified a need for guidance and support to assist entities in achieving a culture of evaluation and learning from experience that underpins evidence-based policy design and delivery.¹⁹

2.2 The Commonwealth Evaluation Policy (the Policy) sets out a principles-based approach to evaluation and applies to all Commonwealth entities. The Policy is supported by Resource Management Guide 130 *Evaluation in the Commonwealth*, also referred to as the Commonwealth Evaluation Toolkit (the Toolkit).²⁰

2.3 It is not feasible, cost effective or appropriate for entities to conduct an evaluation of all government activities and programs.²¹ Hence, entities need guidance on when evaluation is necessary, and what other arrangements should be put in place to ensure the program or policy being implemented is 'evaluation ready'. This includes well-designed data collection and

19 Department of the Prime Minister and Cabinet, *Independent Review of the Australian Public Service*, PM&C, Canberra, 2018, available from <https://www.pmc.gov.au/sites/default/files/resource/download/independent-review-aps.pdf> [accessed 15 January 2024].

20 Australian Centre for Evaluation, *Commonwealth Evaluation Toolkit*, available from <https://evaluation.treasury.gov.au/toolkit/commonwealth-evaluation-toolkit> [accessed 1 April 2024].

21 Australian Centre for Evaluation, *Commonwealth Evaluation Toolkit*, available from <https://evaluation.treasury.gov.au/toolkit/when-evaluate> [accessed 1 April 2024].

performance monitoring arrangements during the design phase of a new or amended program or activity.²²

2.4 The Toolkit states that the evaluation governance actions required to support an evaluation culture include:

- plan to conduct fit-for-purpose monitoring and evaluation activities before beginning any program or activity (identifying timeframes, resources, baseline data, and performance information);
- use strategic risk-based approaches to identify, prioritise, and schedule evaluation activities;
- align internal review, assurance and evaluation activities with external requirements, such as reporting requirements under legislation;
- treat and prepare for external scrutiny as part of standard practice;
- assign responsibility for implementing the recommendations of review, evaluation and performance monitoring activities. Identify who is responsible for implementation and establish timeframes for actions; and
- ensure appropriate monitoring of implementation and the impact on performance and outcomes.

2.5 There is a risk that inadequate evaluation will limit the usefulness of the information collected, contribute to policy or program failure, and inadequately support decision makers in their post-evaluation decisions.

2.6 Appendix 3 contains a summary of the findings from this chapter for Health, DVA and Home Affairs.

Do the entities have a fit-for-purpose approach to choosing when and what to evaluate?

The Evaluation Toolkit provides appropriate high-level guidance to support entities in determining what programs or policies should be evaluated and when. It provides limited guidance on conducting an economic evaluation, including any assessment of cost effectiveness of implementation, and does not include a requirement for all pilots to be evaluated.

- With the exception of DVA's Non-Liability Rehabilitation pilot, at the time the other examined pilots commenced, only Health had established internal evaluation guidance.
- In November 2023, Health published a revised evaluation strategy which specifies roles and responsibilities and includes a tiered system for identifying evaluation priorities across the department.
- Since the commencement of the Wellbeing and Support Program pilot, DVA developed a framework which supports when and what to evaluate based on program characteristics, timing and capability. In August 2023, DVA introduced a framework for the planning, monitoring and evaluation of its health and wellbeing programs, which includes roles and responsibilities.

²² *ibid.*

- Home Affairs does not have an entity-specific approach to determining when and what to evaluate.
- Each entity has an internal evaluation team to provide guidance and support on evaluation practice.

When and what to evaluate

Australian Government guidance

2.7 The Toolkit has principles-based guidance on ‘when’ and ‘what’ to evaluate. The key principles include:

- proportional evaluation effort, to be determined on a case-by-case basis to ensure the overall approach is fit for purpose;
- taking a strategic, risk-based approach to determine when and how evaluations are selected, prioritised and scaled based on the value, impact and risk profile of a particular government activity or program;
- aligning evaluation effort with government or entity priorities, stakeholder priorities, issues with a high profile or sensitivities, and policy or program performance concerns;
- selecting an appropriate evaluation type (formative, summative, rapid);
- selecting an appropriate evaluation approach (depending on the stage of maturity of a program, the issue being investigated, data already available, timing of when findings are needed to support decision making); and
- evaluating at different stages of the policy cycle to support continuous improvement, risk management, accountability and decision making, as needed.

2.8 The Toolkit does not specify any circumstances where an entity must complete an evaluation. There is also no guidance that promotes evaluation of a pilot.

2.9 Pilot programs are not mentioned in the Policy or Toolkit, and are not defined in the following Commonwealth laws, rules or other guidance:

- Public Governance, Performance and Accountability Act 2013* (PGPA Act);
- Public Governance, Performance and Accountability Rule 2014; or
- Budget Process Operational Rules.

2.10 In the absence of a definition within Australian Government guidance, the ANAO used the definition provided by the UN Economic and Social Commission for Western Asia²³, which defines a pilot program as follows:

A pilot program, also called a feasibility study or experimental trial, is a small-scale, short-term experiment that helps an organization learn how a large-scale project might work in practice.

2.11 Pilot programs delivered by Australian Government entities are generally:

- small in scale, whether time-, region- or scope-limited;
- focused on evaluation and learning from experience;

23 United Nations Economic and Social Commission for Western Asia, *Pilot programme*, available from <https://archive.unescwa.org/pilot-programme> [accessed 2 April 2024].

- aimed at testing new approaches; and
- aimed at informing future decision-making.

2.12 The Toolkit includes a wide range of guidance materials from state and territory governments, universities, and overseas entities. No guidance is provided as to which of these materials to follow, or when.

2.13 The Toolkit identifies four categories of questions that can be answered by an evaluation.

- Appropriateness — the extent to which program objectives or desired outcomes align with community needs or government priorities.
- Effectiveness — the extent to which the actual outcomes of a program align to the desired program objectives.
- Cost-effectiveness — the relationship between inputs and actual outcomes, expressed in dollar terms.
- Efficiency — the extent to which inputs are minimised/maximised for a given level of outputs.

2.14 Neither the Policy or Toolkit indicate when evaluations should include an efficiency or cost-effectiveness assessment (an economic evaluation). The Toolkit identifies economic evaluation as a type of evaluation, alongside outcome, impact, summative, post-implementation, and theory of change evaluation. There is no information on how to undertake an economic evaluation or assess whether spending was an appropriate use of public money through the evaluation process.

2.15 The pilots examined during the audit commenced prior to the publication of the Toolkit, with the exception of DVA's NLR pilot, which commenced the month after its publication.²⁴ Prior to the publication of the Toolkit, Australian Government guidance included guidance from the Department of Prime Minister and Cabinet on particular types of evaluation such as regulatory impact analysis²⁵, and the United Kingdom Government's specific evaluation guidance.²⁶

Recommendation no. 1

2.16 The Department of the Treasury update the Commonwealth Evaluation Policy and Toolkit to include:

- a definition of a 'pilot';
- guidance on how to conduct an economic evaluation and other methods for considering whether spending represents an appropriate use of public money;
- a recommendation that evaluations of pilot programs be undertaken;
- a recommendation for evaluation planning to be conducted alongside pilot design; and

24 The audit involved assessing each responsible entity's evaluation frameworks, policies and culture, including alignment with Australian Government guidance and how they were applied to the respective pilots.

25 Department of the Prime Minister and Cabinet, *Australian Government guide to regulatory impact analysis*, 2020, available from <https://nla.gov.au/nla.obj-2825393162/view> [accessed 1 April 2024].

26 HM Treasury, *The Magenta Book: Central Government guidance on evaluation*, 2011, available from <https://www.gov.uk/government/publications/the-magenta-book> [accessed 1 April 2024].

- (e) guidance on governance arrangements for cross-entity evaluations to minimise duplication and maximise coordination and learnings across entities.

Department of the Treasury's response: Agreed.

2.17 *Recommendation 1(a) - Treasury plans to develop a glossary of terms as a future enhancement to the Toolkit, noting definitions will be provided as guidance only (i.e. not mandated) to help build awareness and application of fit-for-purpose evaluation.*

2.18 *Recommendation 1(b) - Treasury plans to enhance the Toolkit to include more practical guidance on analytical methods used in evaluations, including economic evaluation, to improve the way Commonwealth entities assess implementation, measure the impact of government programs, and frame policy decisions on new or significantly amended programs. Future guidance on economic evaluation will complement existing Australian Government resources, such as guidance on cost-benefit analysis from the Office of Impact Analysis, and the Australian Transport Assessment and Planning (ATAP) framework. In relation to broader guidance on whether spending represents an appropriate use of public money, this is already addressed through the suite of guidance administered by the Department of Finance to support Commonwealth entities. The level of detail included in future enhancements to the Toolkit will be designed to complement, and integrate with, that existing guidance.*

2.19 *Recommendation 1(c) - Treasury plans to include further practical guidance in relation to good practice evaluation approaches for pilot programs as part of future enhancements to the Toolkit. The guidance will recommend, but not mandate, that all pilots are subject to formal evaluation. Treasury promotes rigorous evaluation planning for all new policy proposals during the early stages of policy design and implementation to allow the effectiveness and efficiency of programs to be measured and assessed over time, taking into account the Commonwealth Evaluation Principles (fit-for-purpose; useful; robust, ethical and culturally appropriate; credible; and transparent where appropriate). Consistent with the principles-based Commonwealth Evaluation Policy, responsible managers need to determine robust, proportional evaluation approaches for specific pilots or programs delivered in specific contexts. The Commonwealth Evaluation Policy and Toolkit (Resource Management Guide 130) are nested within the public sector resourcing, reporting and performance frameworks established under the Public Governance, Performance and Accountability Act 2013 (PGPA Act). The PGPA framework and supporting guidance, administered by the Department of Finance, establish requirements to support the proper use of public money (i.e. efficient, effective, economical, and ethical) and to ensure Commonwealth entities work cooperatively to achieve common objectives. In this context, the Commonwealth Evaluation Policy provides a principles-based approach for the conduct of evaluations. It highlights the importance of establishing robust monitoring and evaluation arrangements early in the policy cycle, and using strategic, risk-based approaches to plan evaluations. This allows responsible managers to determine proportional evaluation approaches relative to the value, impact, and risk profile of specific programs. The Policy is intentionally principles-based, not prescriptive, does not include mandatory requirements, and does not currently include definitions.*

2.20 *Recommendation 1(d) - Treasury promotes cooperation and effective governance for all evaluations. Further practical guidance in relation to cross-entity evaluations will be considered in the context of future enhancements to the Toolkit. As per the response to recommendation b*

above, the level of detail included in enhanced guidance on the Toolkit will be designed to complement, and integrate with, the broader suite of guidance related to governance and cross-agency cooperation administered by the Department of Finance in support of the Public Governance, Performance and Accountability Act 2013 (PGPA Act) and related whole of government frameworks, such as the APS framework for engagement and participation. Treasury agrees with the recommendation to update guidance in the Commonwealth Evaluation Toolkit. Specifically, Treasury will update the Toolkit to include a definition of a 'pilot', and provide guidance on: economic evaluation, evaluation of pilots, and governance arrangements for cross-entity evaluations. Treasury's guidance on whether spending represents an appropriate use of public money will focus on (and be limited to) guidance on economic evaluation methods, and other fit-for-purpose evaluation approaches. The broader importance of appropriately using public money is well addressed through the suite of guidance administered by the Department of Finance to support resource management and therefore will not be duplicated through Treasury materials. Treasury will recommend, but not mandate, that all pilots are subject to evaluation consistent with the principles-based Commonwealth Evaluation Policy, which recommends that responsible managers need to determine robust, proportional evaluation approaches for specific pilots or programs. Treasury is committed to continuous improvement of the Evaluation Toolkit. Planned enhancement will include more practical guidance on analytical methods, including economic evaluation, and effective governance arrangements that can help to improve the way Commonwealth entities assess implementation, measure the impact of government programs, and frame policy decisions.

Department of Health and Aged Care

2.21 The Department of Health and Aged Care's (Health) Evaluation Strategy 2016–2019 was in effect during the development of the Naloxone pilot (see paragraph 1.9) and Kava pilot (see paragraph 1.15). The strategy contained guidance relevant to all phases of an evaluation. An internal audit of the 2016–2019 Strategy found it had not been fully implemented. The strategy was often not followed and there was a variable level of awareness across the organisation.²⁷ The strategy did not contain any definition of, or guidance relating to, pilot programs.

2.22 A subsequent internal review of the Evaluation Strategy 2016–2019 (which remained in effect until June 2023) resulted in the launch of the Evaluation Strategy 2023–26 in July 2023, which incorporates the evaluation principles from the Toolkit and guidance on ensuring the timing and type of evaluation is fit for purpose. The Evaluation Strategy 2023–26 included four priority areas for action to embed evaluation practices across the department and an implementation plan.

27 The following Auditor-General reports also made findings in relation to poor evaluation practices: Auditor-General Report No.10 2022–23, *Expansion of Telehealth Services*, ANAO, Canberra, 2023, available from <https://www.anao.gov.au/work/performance-audit/expansion-telehealth-services> [accessed 19 March 2024]; Auditor-General Report No.31 2022–23, *Administration of the Community Health and Hospitals Program*, ANAO, Canberra, 2023, available from <https://www.anao.gov.au/work/performance-audit/administration-of-the-community-health-and-hospitals-program> [accessed 19 March 2024]; Auditor-General Report No.8 2023–24, *Design and Early Implementation of Residential Aged Care Reforms*, ANAO, Canberra, 2023, available from <https://www.anao.gov.au/work/performance-audit/design-and-early-implementation-residential-aged-care-reforms> [accessed 19 March 2024]; and Auditor-General Report No.19 2023–24, *Effectiveness of the Department of Health and Aged Care's Performance Management of Primary Health Networks*, ANAO, Canberra, 2023, available from <https://www.anao.gov.au/work/performance-audit/effectiveness-the-department-health-and-aged-cares-performance-management-primary-health-networks> [accessed 19 March 2024].

Action items included Health's internal Evaluation Centre developing a tiering system for identifying evaluation priorities across the department, publishing a rolling schedule of evaluations on an annual basis, identifying valuation requirements in new policy proposals, and ensuring policies and programs are 'evaluation ready'.

Department of Veterans' Affairs

2.23 At the time of commencement of the WASP pilot (see paragraph 1.22) there was no entity-specific evaluation framework within DVA. A draft evaluation framework for pilot programs was prepared in 2017 but not finalised.²⁸ The draft evaluation framework included high-level information about the timing of evaluation activities, such as specifying that a program logic should be developed prior to designing the evaluation, and identified the types of evaluation most relevant to pilot programs as program evaluation and outcome evaluation.

2.24 In July 2020 DVA developed a draft Policy Evaluation Framework that explained the purpose of evaluation, potential objectives of evaluations, and key evaluation principles. The draft Policy Evaluation Framework does not include guidance on what or when to evaluate.

2.25 In May 2022 DVA established a Program Evaluation and Review Plan (the Evaluation and Review Plan), which outlines the schedule of program evaluations and reviews for the health and wellbeing programs to be conducted from 2022–23 to 2024–25. For the 29 programs listed for upcoming evaluation or review, the Evaluation and Review Plan categorises each according to the strategic outcome they relate to and assigns an importance rating of low, medium or high. The NLR pilot is not included in this schedule.

2.26 The Evaluation and Review Plan includes a prioritisation framework which considers program characteristics, timing, evaluation and review type, skills and resource capability and capacity. The prioritisation framework determines the work plan for when evaluation should be undertaken based on consideration of risk, capability, timing, and other factors. There is no definition of what characteristics relate to a 'high', 'medium' or 'low' priority rating. The Evaluation and Review Plan defines each type of review or evaluation activity and uses definitions which align to the taxonomy in the Toolkit.

2.27 The Program Management Framework (PMF) is DVA's overarching guidance for planning, monitoring, and evaluation, and was endorsed in August 2023. The PMF classifies programs as requiring either 'light touch' or 'full scope' documentation based on the size of benefits, materiality of risks, and estimated expenditure of the program. DVA's application of a proportionate and risk-based approach, which is consistent with the Toolkit, means separate monitoring and evaluation plans are not required for 'light touch' programs.

2.28 The PMF indicates the timing of each stage of evaluation, such as specifying that a monitoring and evaluation plan should be created prior to program commencement. Seven templates with accompanying guidance documents are available, covering program logic, monitoring and evaluation plans, program overviews, annual work plan, risk register, lessons learnt, and quality and assurance plans. The PMF does not provide guidance on how to choose an appropriate style of evaluation.

²⁸ Planning documents for the WASP pilot dated January 2017 refer to this draft framework. DVA does not hold a copy of the 2017 document, only a copy of the draft framework dated 12 October 2023.

Department of Home Affairs

2.29 The Department of Home Affairs (Home Affairs) does not have entity-specific rules or guidance for when program evaluation is required, including any entity-specific rules or guidance related to pilots. There is no entity-specific guidance requiring evaluations to be undertaken, no risk-based mechanisms for triggering evaluations, and no guidance on program evaluation.

2.30 In September 2023, the Performance, Governance, and Inquiries Branch at Home Affairs communicated to all Senior Executive Service (SES) officers and Senior Directors a change to its internal approval process for new policy proposals and ministerial and cabinet submissions, stating that:

Performance Governance and Inquiries Branch will support policy teams to consider evaluative thinking and impacts analysis more meaningfully (and earlier on) in the drafting process. This will include ensuring teams are aligning their drafting to the Commonwealth Evaluation Policy Statement and the outcomes aligned to the broader APS reform initiatives related to evaluation.

2.31 The lack of program evaluation guidance at Home Affairs means that there is insufficient structure around decisions to evaluate, and what considerations should be factored into evaluation planning and assessment.

Recommendation no. 2

2.32 The Departments of Health and Aged Care and Veterans' Affairs include in their entity-specific evaluation policies:

- (a) decision-making criteria for the appropriate style of evaluation to be completed by reference to the activity's risk, objective and outcomes;
- (b) guidance on how to demonstrate whether a program represented a proper use of public money, including the cost-effectiveness of its implementation, and how to undertake an economic evaluation where appropriate; and
- (c) guidance related to evaluation of pilot programs.

Department of Health and Aged Care's response: *Agreed.*

2.33 *The Department of Health and Aged Care's Evaluation Strategy 2023-2026 and the new Strategic Investment Framework adopts a risk-based, strategic, and whole-of-department evaluation approach. The evaluation approach is scalable that is proportionate to the size, significance, and risk profile of programs. The department's Evaluation Centre is refining guidance on economic evaluation and evaluation of pilot programs to ensure all programs represent the best value for public money.*

Department of Veterans' Affairs' response: *Agreed.*

2.34 *DVA will include decision-making criteria for the appropriate style of evaluation in the upcoming revision of the DVA Health and Wellbeing Programs Evaluation Plan in 2024. The revised Plan will also include guidance related to the evaluation of pilot programs, in line with the Commonwealth Evaluation Toolkit. DVA's guidance on economic evaluation will be informed by the Australian Centre for Evaluation, and incorporated into DVA's evaluation policies, guidance and materials on an ongoing basis.*

Recommendation no. 3

- 2.35 The Department of Home Affairs develop entity-specific policies for evaluation, including:
- (a) decision-making criteria as to when an evaluation is required and the appropriate style of evaluation by reference to the activity's risk, objective and outcomes;
 - (b) guidance on how to demonstrate whether a program represented a proper use of public money, including the cost-effectiveness of its implementation, and how to undertake an economic evaluation where appropriate; and
 - (c) guidance related to evaluation of pilot programs.

Department of Home Affairs's response: *Agreed.*

2.36 *The department is developing internal guidance material to complement the existing resources available within the Commonwealth Evaluation Policy and Toolkit including the development of a decision-making criteria to inform evaluation activities.*

Accountabilities for evaluation

2.37 The Toolkit sets out the different roles and responsibilities for evaluation.

Department of Health and Aged Care

2.38 The Evaluation Strategy 2023–26 sets out clear accountabilities for evaluation across the department.

2.39 The evaluation plan for the Naloxone pilot identified roles within the department's governance structure for the evaluation, with the exception of not identifying milestones that required sign off by SES officers consistent with the requirements of the 2016–2019 Strategy. An action tracker developed during the pilot implementation stated whether actions were the responsibility of specified departmental staff, the relevant state or territory government or a joint responsibility.

2.40 For the Kava pilot, the evaluation plan was not approved by the relevant SES officer. The evaluation plan does not indicate when advice was due to be provided to the Australian Government, nor outline the broader governance arrangements for the pilot. The Department of Foreign Affairs and Trade undertook a concurrent evaluation of the Kava pilot, focusing on the impact on kava growing countries in the region. There was some overlap between evaluations as the Health evaluation also considered the impact of the pilot on trade opportunities for Australia and Pacific Island countries. There is no record of how the decision was made for both departments to conduct evaluations or how the respective evaluation scopes were determined. The evaluations were undertaken in isolation. The Evaluation Strategy 2023–26 provides no guidance on the governance and accountabilities for cross-entity evaluations, nor is there any guidance in the Toolkit.

Department of Veterans' Affairs

2.41 The PMF clearly outlines accountabilities within DVA for program evaluation. The PMF includes indicative completion times for activities such as the development of the program logic, monitoring and evaluation plan, quality and assurance plans and risk registers.

2.42 DVA's evaluation guidance provides flexibility for who should undertake evaluation, whether they be externally led, policy led or a combination of internal staff and external consultants. The draft Policy Evaluation Framework states that 'equal care and attention is required in commissioning evaluations as in designing an internal evaluation'. The Evaluation and Review Plan states considerations for who will conduct the evaluation include DVA staff capability, program sensitivity, complexity, and available budget.

Department of Home Affairs

2.43 In the absence of entity-specific guidance on evaluations, there are no defined accountabilities and arrangements for evaluations in Home Affairs.²⁹ In the case of the Skilled Refugee Labour Agreement (SRLA) pilot, no evaluation plan was prepared.

Evaluation culture

Department of Health and Aged Care

2.44 The Evaluation Strategy 2023–26 contains four strategic priority areas to promote evaluation culture:

- ensure the most important policies and programs are evaluated;
- planning early for evaluation;
- building capability to conduct and manage high-quality evaluations; and
- enhancing the use of evaluations.

2.45 Supporting actions are identified for the internal evaluation team (the evaluation centre), divisions and executives, along with indicative timelines as part of an implementation strategy. The Evaluation Strategy 2023–2026 requires Health's evaluation centre to report on the performance of the department against the strategy at specified intervals, and to undertake a comprehensive review of the strategy before developing the next one.

2.46 In January 2024, the Executive Committee endorsed the creation of a Chief Evaluation Officer position and the introduction of a pre-registration process for Tier 1 (high-risk) program evaluations, which requires program areas to register the design and analysis plan of an evaluation before it is conducted.

2.47 In January 2024, Health entered into a partnership agreement with the Australian Centre for Evaluation (ACE), in the Department of the Treasury, with the aim to improve the quality and use of evaluations within Health. Health's evaluation centre was consulted in the early stage of the Naloxone pilot, and guidance was provided in relation to the procurement documents and evaluation scope. The evaluation centre was not consulted for the Kava pilot.

29 Aspects of evaluation accountabilities that are not specified include who decides when an evaluation is required, who is accountable for undertaking the evaluation, who has decision-making responsibility for accepting the final evaluation report, and how roles and responsibilities of the entity and contractors are defined for outsourced evaluations.

2.48 The Evaluation Strategy 2023–26 states that the Executive Committee is responsible for deciding whether an evaluation report is published and that the department seeks to progress from its current state of not routinely publishing evaluation reports, to sharing evaluation findings to inform continuous improvement of policies and programs, consistent with a maturing evaluation culture. The Commonwealth Evaluation Principles (see Appendix 4) state that transparency means ‘evaluation findings should be transparent by default unless there are appropriate reasons for not releasing information publicly’. The Evaluation Strategy 2023–26 refers to internal publication of evaluation reports and does not contain guidance on when an evaluation report should be shared publicly.

Department of Veterans’ Affairs

2.49 DVA has a dedicated evaluation team, Program Governance and Evaluation (PG&E), which was established in July 2021. It did not exist during the WASP pilot period. The PG&E team established a Program Management Resource Hub prior to introduction of the PMF, which provides access to resources such as templates and is used to distribute training materials. The PG&E team prepared a training module related to evaluation planning in the PMF. In accordance with PMF principles and requirements, the PG&E team assisted the NLR pilot team in the drafting of the evaluation plan.

2.50 DVA maintains a central repository of audits, reviews and evaluations of health and wellbeing programs with 61 evaluation activities across 31 programs from 2014 to 2023 recorded. The repository does not contain evaluations of programs or activities from other areas across the department. In November 2023 DVA advised the ANAO that the tracker is available to all DVA staff in an unrestricted database and is distributed twice a year to staff with responsibility for program design and delivery. The absence of a repository of evaluation activities across the whole of the department limits the ability to learn from experience and promote evaluation best practice.

2.51 The PMF does not contain a decision-making process for whether evaluation reports or findings are made public. The PMF focuses on internal dissemination of evaluation findings rather than public release, which is inconsistent with the transparency principle in the Toolkit. The draft DVA Policy Evaluation Framework notes that ‘publicly publishing evaluation findings may lead to greater trust from veterans and their families about DVA’s transformation progress, but has political sensitivities’.

Opportunity for improvement

2.52 Health and DVA could amend internal evaluation guidance to be consistent with the Commonwealth Evaluation Principle of transparency, which focusses on public release. Publication of evaluation findings would demonstrate a mature evaluation culture within the organisation.

Department of Home Affairs

2.53 Home Affairs does not have a central repository for evaluations.

2.54 In November 2023 Home Affairs advised the ANAO that its central evaluation team was not consulted for the SRLA pilot.

Recommendation no. 4

2.55 The Departments of Health and Aged Care, Veterans' Affairs and Home Affairs develop and implement explicit guidance to support early engagement with central evaluation teams to improve evaluation strategy and planning.

Department of Health and Aged Care's response: *Agreed.*

2.56 *The Department of Health and Aged Care's Evaluation Centre is refining processes and guidance to ensure that all programs engage early with the Centre and have a fit-for-purpose evaluation plan in place. This approach is consistent with the department's new evaluation strategy.*

Department of Veterans' Affairs' response: *Agreed.*

2.57 *DVA will continue to promote and encourage early engagement to improve evaluation strategy and planning through implementation of the DVA Program Management Framework on an ongoing basis. DVA will continue to develop the Program Management and Evaluation Resource Hub which provides a range of information on effective program management and evaluation and leverages the Commonwealth Evaluation Toolkit and other resources.*

Department of Home Affairs' response: *Agreed.*

2.58 *The department will promote early engagement on evaluation strategy and planning matters, including through publishing updated guidance material.*

Do the entities have an appropriate framework for conducting program evaluations?

Health and DVA have policies and guidance materials for how to conduct program evaluations. Only Health has guidance on when economic evaluation should be undertaken and the guidance is limited. Training on evaluation practices is provided at Health and DVA. Attendance is not consistently monitored. Home Affairs has no entity-specific guidance on conducting evaluations, and no training programs available to staff.

Entity-specific policies and procedures

Department of Health and Aged Care

2.59 The Evaluation Strategy 2023–26 incorporates guidance from the Commonwealth Evaluation Policy and Toolkit and relevant evaluation features and effort required, depending on policy or program characteristics. The Evaluation Strategy 2023–26 lists 'process', 'outcome' and 'economic' as the three main forms of evaluation and the factors that are relevant in determining which method is appropriate.

2.60 Specific guidance on how to conduct an evaluation is not included in the Evaluation Strategy 2023–26. No guidance is included in the Evaluation Strategy 2023–26 as to when an economic evaluation should be undertaken, or how else to demonstrate that the implementation of the policy or program was an appropriate use of public money.

Department of Veterans' Affairs

2.61 The Program Management Framework (PMF) provides clear guidance on the steps required to plan for and undertake an evaluation, including a summary of the deliverables that should be prepared throughout the process (e.g. program logic) and a description of the support available from the PG&E team. The PMF does not provide guidance on when an economic evaluation should be undertaken, or how else to demonstrate that the implementation of the policy or program was an appropriate use of public money.

2.62 The PMF sits within a broader suite of resources designed to provide staff with guidance on program monitoring and evaluation. Other DVA frameworks include the Policy Evaluation Framework, Provider Engagement Framework, Risk Management Framework, Assurance Framework and Outcomes and Benefit Realisation Framework.

2.63 As at March 2024 DVA's specific guidance for the evaluation of pilot programs is in draft form. The draft Evaluation Framework for Pilot Programs states that it 'aims to provide a consistent approach to evaluating pilot programs across DVA, in particular pilot programs which are part of Veterans Centric Reform'. The draft Evaluation Framework for Pilot Programs provides high-level guidance on each step of evaluating pilot programs, including the importance of baseline data, and provides a series of templates, such as an example set of evaluation questions and a data collection strategy.

Department of Home Affairs

2.64 Home Affairs does not have entity-specific rules or guidance for how to conduct program evaluation (see paragraph 2.35 for related recommendation).

Opportunity for improvement

2.65 The Commonwealth Evaluation Policy and Toolkit states that data analysis methods are a key consideration in assessing whether program results met intended objectives and provides guidance on how to answer causal questions. Entity-specific guidance could cover approaches to data analysis and answering causal questions, seeking to go beyond the generic guidance in the Toolkit and include entity-specific considerations. This could include an overview of available datasets, advantages and disadvantages of different approaches, and examples of previous evaluations by the entity that applied best-practice approaches.

Training about how to evaluate

Department of Health and Aged Care

2.66 Health offers its staff evaluation related training via workshops and webinars, including workshops on program logic, indicator development and the Monitoring and Evaluation Framework.

2.67 The evaluation and training landing page on Health's intranet includes links to publicly available videos on evaluation from a range of external sources and covers various aspects, such as developing key performance indicators, measuring outcomes, ethics, and budgeting. The linked materials are not specifically related to Health's Evaluation Strategy 2023–26, but broadly reflect evaluation best practice. The evaluation centre recorded that 219 staff attended program logic

workshops to 'support evaluation procurement' between September 2022 and January 2024. There are no testing protocols to ensure staff understand the material at the conclusion of the training.

Department of Veterans' Affairs

2.68 The PG&E team provides on-the-job training to DVA staff through presentations and feedback on draft evaluation related documents. In May 2023, the PG&E team facilitated a training module for the Client Engagement and Support Services Division on evaluation planning in the PMF. In November 2023, DVA advised the ANAO that this training was mandatory for Client Engagement and Support Services Division staff. There is no evidence of monitoring or reporting of compliance with this requirement. In June 2023, the PG&E team delivered a presentation to DVA staff about the implementation of the PMF and key activities that contribute to a well-initiated and planned project. The presentation was scheduled to run for ten minutes, and a majority of divisions were represented among the 118 attendees.

Department of Home Affairs

2.69 Home Affairs does not provide any evaluation training. The central evaluation team supports conduct of evaluations by encouraging use of Toolkit templates and providing feedback on evaluation materials such as plans and reports.

3. Evaluation approach

Areas examined

This chapter examines the appropriateness of evaluation planning for the selected pilots, including whether the evaluation methodologies were consistent with good practice.

Conclusion

The evaluation planning and approach for the Department of Health and Aged Care's (Health) Take Home Naloxone pilot and the Department of Veterans' Affairs' (DVA) Wellbeing and Support Program pilot were largely robust, including appropriate stakeholder engagement and relevant ethics approvals. Planning for the evaluation of Health's Kava pilot did not identify the risk that that ethics approval may not be granted for one of the planned qualitative analysis methods, and there was a lack of baseline evidence to support the planned evaluation methodology. The effectiveness of planning for the evaluation of DVA's Non-Liability Rehabilitation pilot was reduced as the analytical methodologies were not documented, and no external stakeholders were consulted. The Department of Home Affairs did not complete its planning for, or undertake, a robust evaluation for the Skilled Refugee Labour Agreement pilot. All evaluation plans and approaches could have been enhanced by a greater focus on the availability of data and an assessment of the proper use of public money.

Areas for improvement

The ANAO made one recommendation in relation to the approval and content of evaluation plans.

3.1 The Resource Management Guide 130 *Evaluation in the Commonwealth*, also referred to as the Commonwealth Evaluation Toolkit (the Toolkit)³⁰ states that:

It is good practice to

- plan how government activities and programs will be monitored and evaluated from the start [and]
- collect robust performance information that supports continuous improvement, risk management, accountability, and decision-making at all policy stages.

3.2 The evaluation principles within the Toolkit are set out in Appendix 4.

3.3 The Toolkit sets out what needs to be considered to ensure that a fit-for-purpose evaluative approach is developed. This includes:

- before an evaluation begins, evaluators should investigate the context, set evaluation objectives, determine the scope and approach, and define evidence and data sources;
- throughout an evaluation, evidence and data should be collected; and
- early consideration should be given to how to analyse and interpret results, report findings, and implement improvements, to ensure stakeholders get what they need from the evaluation.

30 Australian Centre for Evaluation, *Commonwealth Evaluation Toolkit*, available from <https://evaluation.treasury.gov.au/toolkit/commonwealth-evaluation-toolkit> [accessed 1 April 2024].

3.4 Poor planning processes can contribute to missed opportunities to identify and capture evidence at critical intervals, particularly in relation to data to support a comparator to an intervention, which can impede the quality of evaluation. It also creates a risk that the evaluation will not provide the evidence needed to support decisions about post-pilot policy continuance or amendment, potentially voiding the value of, and investment in, the pilot and its evaluation.

3.5 Appendix 5 contains a summary of the findings from this chapter for each of the pilots examined during the audit.

Was there appropriate evaluation planning before beginning the pilot?

Planning for evaluation, including stakeholder engagement, was completed for Health's Take Home Naloxone pilot, and DVA's Wellbeing and Support Program pilot. Planning for stakeholder engagement for the evaluation of Health's Kava pilot did not account for the risk that ethics approval may not be granted and the resulting impact on the planned analysis and evaluation methodology. The effectiveness of planning for the evaluation of DVA's Non-Liability Rehabilitation pilot was reduced as the methodologies to be used were not documented and no external stakeholders were consulted. Home Affairs did not complete its planning for the evaluation for the Skilled Refugee Labour Agreement pilot. While data sources were identified within the evaluation plans that were developed, one or more planned data sources within each pilot were not available for the evaluation and this risk had not been identified.

Fit-for-purpose evaluation plans

3.6 In describing 'how to evaluate', the Toolkit categorises evaluative steps into three categories: planning and budgeting, measuring and assessing, and reporting and accountability. The 'planning and budgeting' step has three components:

- investigate context;
- set evaluation objectives; and
- determine scope and approach.

3.7 The Toolkit also states 'define evidence and data sources' needs to be considered before an evaluation begins. Early planning about what the evaluation should achieve or deliver will likely lead to a more robust and useful evaluation report.³¹ Effective early stakeholder engagement is also important to provide various perspectives on what will be a useful and credible evaluation, contributing to the program logic, and facilitating quality data collection.

3.8 The Toolkit states that before a program is implemented, entities need to plan how a program will be evaluated.

Department of Health and Aged Care

Take Home Naloxone pilot

3.9 The evaluation plan for the Department of Health and Aged Care's (Health's) Take Home Naloxone pilot (the Naloxone pilot) (see paragraph 1.9) was largely fit for purpose and covered all phases of the evaluation including governance arrangements, an evaluation framework, program

31 Australian Centre for Evaluation, *Commonwealth Evaluation Toolkit*, available from <https://evaluation.treasury.gov.au/toolkit/what-evaluation> [accessed 1 April 2024].

logic, ethics strategy, and risk management plan. The robustness of the evaluation plan was reduced by the absence of an economic evaluation or any other assessment of the cost effectiveness of delivering the pilot's objective.

3.10 The Toolkit highlights the importance of creating counterfactuals to use to compare against the results of pilot participants. A case study approach was selected to assess the effectiveness of the Naloxone pilot. Impacts on participant access and uptake were based on interviews during site visits at three intervals with target populations, their peers, and families. Appropriate planning for a counterfactual would have increased the effectiveness of the trial and conclusions drawn from the evaluation.

3.11 The evaluation's quantitative data analysis assessed trends in naloxone supply through the Pharmaceutical Benefit Scheme (the PBS) for pilot and non-pilot jurisdictions, and rates of accidental opioid overdose in the pilot jurisdictions only. The analysis was consistent with the evaluation plan. Volumes supplied through the PBS were largely unaffected over the duration of the pilot, in both pilot and non-pilot jurisdictions, which limited the explanatory power of any pilot versus non-pilot jurisdictional comparison. There is a risk that the exclusion of non-pilot states from the evaluation scope limited the analysis to be undertaken during the evaluation.

Kava pilot

3.12 For the Kava pilot (see paragraph 1.15) the evaluation plan covered relevant aspects of planning: evaluation questions; evaluation design (program logic and methods); community and stakeholder engagement; key tasks and milestones; risk assessment and management plan; and governance arrangements for the evaluation activities.

3.13 The robustness of the evaluation plan was reduced as:

- evaluation questions were repetitive, and, in some cases, there was significant overlap;
- it did not reflect the evaluation being undertaken by the Department of Foreign Affairs and Trade and seek to avoid duplication or inconsistencies;
- it did not identify the approach to establish baseline kava use, or establish how long-term impacts of kava use could be measured during the two-year evaluation period;
- it did not anticipate that ethics approval may not be obtained, which prevented the evaluation from consulting with members of Aboriginal and Torres Strait Islander communities in the Northern Territory who were identified as a high-risk group for potential harmful kava impacts;
- the risk of stakeholders not participating in the evaluation was not identified or mitigated in the risk assessment and management plan (see paragraph 3.34); and
- the logic map was not consistent with internal guidance and did not map the inputs, activities and outputs to the specific outcomes measured in the evaluation report.

3.14 The evaluation plan did not consider whether the evaluation would be able to answer questions about harms and availability of evidence to support a conclusion on this objective. This risk was not identified in the risk assessment and management plan. In November 2023 Health advised the ANAO that planning for the Kava pilot had not commenced prior to the pilot's announcement by the Prime Minister.

Department of Veterans' Affairs

Wellbeing and Support Program pilot

3.15 The Wellbeing and Support Program (WASP) pilot's draft evaluation plan was largely fit for purpose and covered all phases of the evaluation, including a timeline, draft program logic, implementation analysis framework, data collection plan, relevant stakeholders and qualitative and quantitative analysis methodology.³² The draft evaluation plan defined six overarching research questions, none of which related to the cost-effectiveness of the pilot.

3.16 The plan identified risks, such as difficulty following up with participants after the pilot and using self-reported client data, with no management plan developed. The draft evaluation plan did not identify other risks such as selection bias among pilot participants, missing or erroneous data, and pre-pilot quality of life measures. Alternative methodological approaches such as experimental approaches or quasi-experimental approaches using a control group were not considered (see paragraph 3.64).

3.17 The funding proposal for the evaluation was approved by the DVA Research Board in accordance with departmental policy requirements. The contracted evaluation team was involved in evaluation planning and was engaged three months after the pilot commenced.

Non-Liability Rehabilitation program pilot

3.18 The evaluation plan for the NLR pilot is undated. In November 2023, DVA advised the ANAO that the evaluation plan is being updated following the introduction of the Program Management Framework (PMF) (see paragraph 2.27).

3.19 The plan states that the evaluation will be conducted in two stages: a process and implementation evaluation (the mid-pilot review) and a post-implementation review assessing the appropriateness, efficiency and effectiveness of the pilot, including identifying lessons learned and potential future improvements. The mid-pilot review is aimed at assessing whether the pilot is working as expected, if adjustments should be made to design or delivery, and if there are unintended consequences or gaps.

3.20 The evaluation plan includes evaluation objectives, available data, relevant stakeholders, the evaluation questions for both the mid-pilot review and post-implementation review. The evaluation questions for the post-implementation review are themed into the following headings: 'appropriateness, efficiency, and effectiveness and impact'. The program logic has the same structure as the PMF template. An evaluation framework was developed that maps the evaluation questions to themes (such as effectiveness), measures and data sources, and method of collection. The evaluation plan lacks a methodology for analysis of the data it identifies as relevant to the evaluation questions. It states:

The evaluation will also compare the outcomes achieved by NLR participants with those who undertake standard rehabilitation (compulsory rehabilitation as well as voluntary). However, there will be limitations in this comparison due to differences in intent and design of the programs e.g. services offering, financial limits etc.

3.21 The absence of a quasi-experimental approach limits the effectiveness of the evaluation plan, as it is unlikely the evaluation would be able to conclude against the causal impact of the pilot,

³² DVA advised the ANAO that the evaluation plan was finalised and endorsed. The plan is labelled draft on the front cover.

such as the question ‘did NLR result in positive changes to participants’ social wellbeing and employment?’. The evaluation plan includes a question about whether the NLR program offers value for money, but no method for economic evaluation is specified. The evaluation plan also lacks consideration of risks and limitations.

Department of Home Affairs

3.22 Home Affairs undertook preliminary planning but did not complete an evaluation plan for the SRLA pilot. A draft evaluation ‘skeleton’ was prepared, which outlined the structure of an evaluation report and included sections for the program purpose, audience, scope, methods, risks, and recommendations of the evaluation. Home Affairs undertook one of the four steps recommended in the Toolkit (investigate context). It did not complete the other three steps (set evaluation objectives, determine scope and approach, and define evidence and data sources).

3.23 During the planning for the pilot, Home Affairs advised the minister in October 2019 that:
subject to first year outcomes, the Pilot could then be expanded and a program for skilled refugees established, taking into consideration required legislative and system changes.

3.24 There was no evidence of Home Affairs defining the outcomes expected from the first year of the pilot, other than the number of visa applications that would be lodged and granted, or how such outcomes would be measured or evaluated.³³ There was no evidence of the decision-making process as to the type of evaluation required. No governance arrangements for the evaluation were established, the method of evaluation or evaluation questions were not defined, no program logic was developed, and the time and resources required for an evaluation were not considered.

Recommendation no. 5

3.25 The Departments of Health and Aged Care, Veterans’ Affairs and Home Affairs ensure evaluation plans are prepared for policies or programs subject to evaluation requirements and that they be approved prior to the implementation of the policy or program. Consistent with the Commonwealth Evaluation Toolkit, evaluation plans should incorporate proportionate and risk-based level of information, including:

- (a) methods for measuring or capturing baseline evidence, and attributing changes to the pilot, policy or program; and
- (b) a method of economic evaluation or other means of assessing the proper use of public money.

Department of Health and Aged Care’s response: Agreed.

3.26 *The Department of Health and Aged Care is formalising processes to ensure high risk programs have a fit-for-purpose evaluation plan in place during program development. The evaluation plan includes a) program’s logic model; b) key evaluation questions/ indicators/ data sources and analytical approaches and c) appropriate types and timing of future evaluation. This approach is consistent with the Department’s new evaluation strategy and the Commonwealth Evaluation Toolkit.*

33 Between 1 July 2021 to 23 May 2023, 73 visa applications were lodged under the SRLA pilot. All applications were for Employer Nomination Scheme (subclass 186) visas with 46 visas granted to primary applicants plus 72 dependent applicants.

Department of Veterans' Affairs' response: Agreed.

3.27 *DVA will continue implementing the Program Management Framework to allow for a proportionate and strategic risk-based approach to evaluation, including economic evaluation where appropriate. DVA will establish an enterprise-level oversight function using an existing governance body to approve evaluation plans for all policies, programs, and pilots subject to evaluation requirements, prior to implementation.*

Department of Home Affairs' response: Agreed.

3.28 *The department agrees that evaluation plans (that meet the decision-making criteria referred in recommendation No.1) should be approved prior to the implementation of the policy or program wherever possible, including the information detailed in the recommendation (paragraph 3.25 (a) and (b)). The department's central evaluation function will assist in the co-design of evaluation artefacts during the policy and program development stage. The SRLA Pilot reviewed in this report is already operational. Evaluation requirements for the SRLA Pilot are being developed and an evaluation plan and supporting program logics are being developed.*

Stakeholder engagement

3.29 The Toolkit provides guidance about the role of stakeholder engagement in evaluation. The Toolkit states that stakeholders should be identified to ensure the evaluation meets evaluation principles of usefulness, transparency, credibility, and robustness, and is fit for purpose.

3.30 The Toolkit recommends that during the planning and budgeting stage, key stakeholders for an evaluation are identified and input sought on evaluation objectives, questions, data collection methodologies, scope, and approach.

*Department of Health and Aged Care**Naloxone pilot*

3.31 The pilot's project plan identified external stakeholders to be involved in the development of the pilot and its evaluation plan. A stakeholder roundtable was held to inform the pilot's design, and the program logic captured issues across areas of stakeholder interest. The stakeholder communications log identified the key meetings held and the written correspondence sent. The information captured in the log is consistent with the pilot's communications plan that outlined the types of communications and target audiences.

Kava pilot

3.32 A stakeholder consultation paper was prepared for each phase of the pilot. Other than the collation of the feedback received on the consultation papers, the outcomes of consultations with the Western Australian Government were the only consultations with state and territory governments documented. In November 2023, Health advised the ANAO that 'the pilot was implemented as requested by Government and the design was not directly influenced by any specific external consultation process undertaken'.

3.33 The evaluation report stated:

there was disappointment at lack of consultation with Aboriginal and/or Torres Strait people on the possible impact the change to kava laws may have on ... communities and individuals ... overall impressions highlight a lack of satisfaction with the consultation process.

3.34 As part of the evaluation, surveys, interviews and focus groups were undertaken to collect information from stakeholders. The risk of stakeholders not participating in the evaluation was not identified or mitigated in the evaluation plan risk assessment or management plan.

- The community survey was only administered for Pacific Islander communities, no response was received from the Australian South Sea Islander organisations invited to participate in the data collection.
- Consultations did not occur with members of Aboriginal and Torres Strait Islander communities in the Northern Territory as ethics approval was not granted (see paragraph 3.13).

Department of Veterans' Affairs

Wellbeing and Support Program (WASP) pilot

3.35 DVA reported internally that workshops were undertaken with the Department of Defence (Defence) in the design of the pilot. There was no evidence of other stakeholder consultation in the design of the pilot, prior to the design of the evaluation activities.

3.36 Consultation methods for the evaluation included interviews, surveys, focus groups, and email feedback. Stakeholders consulted as part of the qualitative analysis included a subset of veterans, their families, case managers, health and psychosocial service providers, and DVA staff. Stakeholder engagement was conducted according to the plan, and formal data collection approaches were used.

Non-Liability Rehabilitation (NLR) program pilot

3.37 DVA consulted internally on the design of the pilot, particularly in relation to managing legal risks. A communications strategy was developed for the implementation of the pilot. The strategy refers to a mix of communication products to be used for the key target audiences, including a planned approach to minimise risk of confusion amongst the veteran community as to the purpose of the pilot.

3.38 The evaluation plan for the NLR pilot identified relevant stakeholders for the report itself, not how these stakeholders would be involved in the evaluation exercise or their particular views or needs. The mid-pilot review states that it involved discussions with pilot stakeholders including departmental staff, pilot participants and rehabilitation service providers. The process and findings of this engagement are not described in the report. The mid-pilot review assessed the effectiveness of the communications strategy with external and internal stakeholders undertaken to recruit pilot participants, with low pilot take-up being attributed in part to confusion with NLR language and messaging.

Department of Home Affairs

3.39 In the development of the SRLA pilot, Home Affairs advised the Minister for Immigration, Citizenship, Migrant Services and Multicultural Affairs that various branches within the department were consulted, and external consultation was only with Talent Beyond Boundaries (TBB).³⁴

34 Home Affairs advised the Minister that Talent Beyond Boundaries (TBB) was 'an Australian and international not-for-profit organisation specialising in finding migration pathways for skilled migrants and refugees'. TBB provided advice, information, and some activities in support of Home Affairs delivering the pilot.

3.40 In July 2023, Home Affairs advised the Minister for Immigration, Citizenship, Migrant Services and Multicultural Affairs, that the SRLA pilot:

has received positive feedback from sponsoring businesses and candidates and is highly regarded as best practice and proof of concept by international stakeholders and likeminded countries looking to implement similar programs.

3.41 There was no evidence of how Home Affairs collated or verified this feedback or the benchmarking undertaken in relation to similar programs, which is inconsistent with the principles of credibility and robustness in the Toolkit.

Identification of data sources

3.42 The Toolkit states that defining evidence and data sources should occur before an evaluation begins, and recommends consideration of:

- the best measures to use in the evaluation;
- having 'a plan for what to measure, when and how to collect data and from whom'; and
- how to design or source data/evidence collection tools.

Department of Health and Aged Care

Naloxone pilot

3.43 The quantitative data collection strategy set out in the evaluation plan included:

- data from the point of naloxone supply (date, postcode, site type, voucher type (relevant for South Australia), naloxone formulation and quantities, consent statement, initial supply or refill choice, reason for refill, participant type, source of opioids used, other medications or substances in use, education or advice provided, resources provided, gender, age, staff designation);
- bulk supply arrangements via wholesale arrangements;
- data to monitor the trajectory of the Naloxone pilot indicators, including PBS dispensing data, illicit drug reporting survey data, hospital emergency department data, ambulance data, and drug-related deaths data; and
- accidental deaths from opioid overdose from emergency departments or ambulance services.

3.44 Despite being referenced in the evaluation plan, there was no data on accidental deaths from opioid overdose from emergency departments or ambulance services available for analysis during the evaluation.

3.45 The evaluation plan also referred to evidence to be collected from round table discussions with stakeholders, and site assessments and interviews at 12 pilot sites. The sampling approach, data collection, and questions for the stakeholder roundtables and the recruitment strategy for interviews and data collection for the site assessments were all defined.

Kava pilot

3.46 Strategies for data collection for the Kava pilot were outlined in the evaluation plan. The strategies included community surveys, semi-structured interviews, Nominal Group Technique and Yarning Circle focus groups³⁵, 'routinely collected' national data, and a systematic literature search.

3.47 'Routinely collected' data used for the evaluation included annual data from face-to-face interviews with drug users (Illicit Drug Reporting System and Ecstasy and Related Drugs Reporting System), data from drug treatment services (Alcohol and Other Drug Treatment Services National Minimum Data Set), death investigation records provided by the coroners' courts (National Coronial Information System) and from data on the monthly number of kava permits issued, and volume of kava imported since 1 December 2021 to 31 May 2023 from the Office of Drug Control. The NSW Ministry of Health also shared data with the evaluation team.

Department of Veterans' Affairs

Wellbeing and Support Program (WASP) pilot

3.48 A data collection plan was included in the draft evaluation plan, detailing the type and quantity of interviews required for the qualitative analysis. Data was planned to be de-identified for privacy and anonymity. A recruitment strategy was developed for clients, carers, case managers and a focus group, with multiple options to account for varying logistical concerns, access, and levels of burden on those supporting recruitment. The data collection strategy covered baseline data, data from during the pilot period, and post-pilot follow up.

3.49 The quantitative analysis method considered the balance between collecting sufficient data and minimising burden on data collectors. For example, data about quality of life, healthcare utilisation and hospital admissions could be collected from the client and case manager, but this was considered by the evaluation team to create a significant record-keeping burden. Health utilisation data (e.g. the number and type of DVA and hospital encounters before and after the pilot by clients) was planned to be directly sourced from DVA to minimise the burden on participants. A range of validated clinical measures and health utilisation data were identified for quantitative analysis, with details on how the quantitative analysis could be undertaken.

3.50 The draft evaluation plan specified that minimum measures for all clients included three different quality-of-life or wellbeing measures, specifically the SF12³⁶, Partners in Health³⁷, and Personal Wellbeing Index³⁸ measures. There was limited consideration of the risk of gathering insufficient data about participants during and after pilot participation (see paragraph 3.16). There was no consideration of the risk of selection bias.

Non-Liability Rehabilitation (NLR) program pilot

3.51 The evaluation plan defines a set of key indicators to address the evaluation questions, with a corresponding data source and data collection method for the pilot period. For example, one

35 Nominal Group Technique is a structured consensus-based method of collecting information. Yarning Circle is an in-depth discussion undertaken in a culturally respectful environment. See National Drug and Alcohol Research Centre, *Monitoring and Evaluation of the Kava Pilot Program*, report prepared by Asare-Doku W et al., NDARC, 2023, p. 26, available from <https://www.odc.gov.au/sites/default/files/2023-12/monitoring-evaluation-kava-pilot-program.pdf> [accessed 12 June 2024].

36 SF-12 is a 12-item general health questionnaire.

37 The Partners in Health tool is a questionnaire to assess self-care behaviours.

38 The Personal Wellbeing Index is an index designed to measure subjective wellbeing.

evaluation question is whether rehabilitation uncoupled from compensation improves patient outcomes. To answer this question, the evaluation plan identifies the outcomes to be measured as achievement of goals and improvement in social wellbeing and employment status, collected via self-assessed closure feedback forms. An assessment of each data source's reliability and accuracy is also included. In addition, the evaluation plan identifies what existing data sources can be leveraged for the evaluation, including comparator data.

Department of Home Affairs

3.52 During the planning phase of the SRLA pilot in 2020, Home Affairs advised the Minister for Immigration, Citizenship, Migrant Services and Multicultural Affairs that 'annual reporting will also be undertaken to ensure outcomes are consistent with the policy intent of the Pilot and the requirements set under the labour agreement'. Home Affairs developed a spreadsheet containing the characteristics of each pilot participant and the nominating employer, the status of their visa, and measures of processing time. There was no evidence of any other plan to collect data relating to the pilot to support the evaluation.

Are evaluation criteria methodologies consistent with good practice?

The evaluation methodologies used for three out of the five pilots examined were largely consistent with the Toolkit. For the evaluations conducted, all could have been strengthened with a greater focus on baseline data, control group outcomes, and an assessment of the proper use of public money. Ethics approvals were obtained for Health's Naloxone pilot and DVA's WASP pilot. The ethics approval sought for Health's Kava pilot was not granted and no alternative strategy was developed to obtain information that was critical to the evaluation. DVA's Non-Liability Rehabilitation pilot evaluation plan did not include a consideration of ethics approval and the post-implementation review has not yet been undertaken. As Home Affairs did not conduct an evaluation of its pilot there was no methodology applied, or consideration of the need for ethics approval.

Robustness of methodologies

Department of Health and Aged Care

Naloxone pilot

3.53 The evaluation methodology for the Naloxone pilot was consistent with the Toolkit and evaluation plan. The deviations from the evaluation plan included the limited focus on providing access to naloxone beyond recognised sites within the alcohol and other drugs sector, and no public launch or awareness campaign to promote Take Home Naloxone program to the public.

3.54 Limitations in the evaluation methodology are set out below.

- There was risk that using pre-trial naloxone supplies in pilot sites as the base case for the purposes of the evaluation would ascribe any change in naloxone supplies to other factors, such as an increase in opioid use, particularly given the pilot occurred during the COVID-19 pandemic. This risk was identified in the evaluation report and was partly mitigated by the comparison in the final evaluation report between naloxone supplied through the Naloxone pilot and that supplied through the PBS (in all jurisdictions). An attempt was also made to account for external factors that may have impacted on PBS dispensing patterns

in the pilot and non-pilot jurisdictions. This made, what was otherwise relatively limited, quantitative comparisons between pilot and non-pilot jurisdictions more robust.

- The base case selected for the evaluation was predominantly a 'before-and-after' approach, examining trends in supplies of naloxone over the course of the trial, and interviewing target populations, their peers, and families, during site visits at three time-points to ascertain enablers, barriers and impacts over time. This assessed the experience of those in the pilot relative to what they might otherwise have experienced, rather than comparing outcomes with those in non-pilot participating jurisdictions. All the research questions for the evaluation relied on the before-and-after method of selecting the base case, supplemented by comparisons in naloxone activity in pilot and non-pilot jurisdictions. This risk was not identified in the evaluation report.
- No economic evaluation was undertaken or other analysis of the delivery of the pilot which would inform future program design, including whether the implementation of the pilot was a cost-effective way of delivering the pilot, or whether the pilot represented an appropriate use of public money.

Kava pilot

3.55 The evaluation methodology for the Kava pilot consisted of mixed methods research design, including quantitative data (cost-benefit analysis, routinely collected data, community surveys) and qualitative data (for example, interviews and focus groups with government and non-government stakeholders). The methods used were partly robust and largely consistent with the Toolkit.

3.56 Limitations in the evaluation methodology are set out below.

- The evaluation did not define the base case or how much kava would have been imported or consumed without the pilot. The timing of the evaluation limited the ability to assess kava use under the two-kilogram personal import limit, or prior to commercial importation.
- The cost-benefit analysis concluded that benefits of the pilot were significantly larger than the costs or harms of kava use. This was not consistent with findings from the literature review, which found harms of kava use are uncertain.
- Several costs and benefits measured in the cost-benefit analysis are based on community survey data (see paragraph 3.34) which is not generalisable and may be biased.
- A discrete-choice experiment (DCE)³⁹ was used to estimate the value of increased access to kava. This method was used to monetise the social and cultural benefits to Pacific Islander communities in Australia, the strengthened diplomatic relationships, and the economic benefits to Pacific Island countries. The evaluation report noted that:
 - the survey was only administered to government stakeholders (i.e. only government stakeholder preferences were used to value user benefits) which are unlikely to be representative of kava users or the broader community;

39 A discrete choice experiment is a quantitative methodology used to elicit preferences from participants. A DCE involves presenting alternative hypothetical scenarios for participants to choose from, each with different attributes. Participant choices can be analysed to infer their preferences about these attributes. See York Health Economics Consortium, *Discrete Choice Experiment (DCE)*, YHEC, 2016, available from <https://yhec.co.uk/glossary/discrete-choice-experiment-dce/> [accessed 12 June 2024].

- survey sample size was 28 participants. A larger sample would have seen more statistically significant attribute parameter estimates.

3.57 The valuation of user benefits via the DCE was not appropriate due to the survey sample size, choice of respondents and that there already was a market for kava use which could have been used to estimate user benefit.

3.58 While the evaluation collected qualitative information on impacts of kava use and stakeholder views of the pilot, the data collected was not sufficient to provide conclusive evidence about harms and benefits of kava use. The evaluation report recommended ongoing monitoring of kava use and resulting harms, which was supported by stakeholders.

Department of Veterans' Affairs

Wellbeing and Support Program (WASP) pilot

3.59 The WASP pilot evaluation methodology was partly robust, involving a mixed methods approach to undertake a process evaluation and outcomes evaluation.

3.60 The process evaluation of the WASP pilot was undertaken by applying the Realist Evaluation⁴⁰ and RE-AIM⁴¹ methodologies. The Realist Evaluation framework is an effective method to evaluate health and social care programs. The RE-AIM methodology is a framework to understand the relative strengths and weaknesses of different approaches to health promotion and chronic disease self-management.

3.61 The evaluation plan outlined that economic analysis would be undertaken, including a methodology for measuring changes in health services utilisation, health services costs, quality adjusted life years, and wellbeing and psychological functioning. The methodology does not take these estimates and use them to conduct a cost-benefit analysis or cost-effectiveness analysis.

3.62 The quantitative methodology in the draft evaluation plan did not define a control group. The planned methodology for analysis of healthcare utilisation data involved an interrupted time series analysis, which is a robust quasi-experimental approach where a counterfactual can be established without a control group. This approach proposed to use monthly measures of cohort specific measures (such as hospitalisation rates). The draft evaluation plan does not include consideration of undertaking a randomised controlled trial or otherwise defining a control group. The plan states that a generalised linear mixed model framework would be employed in the analysis of quality-of-life and a repeated measures Analysis of Variance (ANOVA) approach will be used to assess statistical significance of changes in measures of wellbeing and psychological functioning.

3.63 The methodology applied in the evaluation was partially consistent with the planned methodology. Analysis of healthcare utilisation data did not follow the planned quasi-experimental

40 Realist Evaluation applies a context-mechanism-outcome structure to understand the mechanisms by which the intervention results in changes, and the impact of contextual factors on those mechanisms. See Pawson, R. and Tilley, N. *Realistic Evaluation*, 2024, available from <https://www.urban-response.org/system/files/content/resource/files/main/pawson---tilley-%282004%29-realistic-evaluation.pdf> [accessed 1 April 2024].

41 RE-AIM determines the population-based impact of a program or policy by assessing interventions across five dimensions: Reach, Effectiveness, Adoption, Implementation and Maintenance. See Glasgow, R.E., Vogt, T.M. and Boles, S.M. *Evaluating the public health impact of health promotion interventions: the RE-AIM framework*, American Journal of Public Health, 1999 September, volume 90, September 1999, pp 1322-1327, available from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1508772/> [accessed 1 April 2024].

(interrupted time series) approach, with a simpler linear mixed model approach being used. This approach involved comparing costs during periods before WASP against costs after entry to WASP. Such pre-post analyses provide a lower standard of evidence than quasi-experimental approaches, because they assume that there are no other factors affecting outcomes over time.⁴²

3.64 The planned data collection for the evaluation had a risk of selection bias from multiple sources. One form of selection bias noted in the limitations section of the evaluation report is where those volunteering to participate in the pilot were those with greater approval levels of WASP compared to non-participants. Only some participants had pre-post data collected, and those who provided data may be systematically different from those who did not. The WASP evaluation methodology relied substantially on survey data for the qualitative analysis and parts of the quantitative analysis, and neither risks of selection bias nor experimental approaches were considered in the plan (see paragraph 3.16). Quasi-experimental approaches involving a control group such as regression discontinuity⁴³ or difference-in-differences⁴⁴ were not considered in the plan, with both being noted as robust approaches in the Toolkit.

3.65 Analysis of healthcare utilisation included costs incurred by DVA and not any costs borne by other areas of government, which is identified as a limitation in the evaluation report (see paragraph 4.13). This is contrary to the evaluation plan statement that the evaluation will ‘examine the effectiveness of the Pilot from the perspective of all stakeholders’.

Non-Liability Rehabilitation (NLR) program pilot

3.66 The NLR evaluation plan does not contain an analysis methodology (see paragraph 3.20). The evaluation plan identifies data for a comparator group that undertakes standard rehabilitation (compulsory rehabilitation as well as voluntary). The approach to analysing this data to support findings is not identified in the plan.

3.67 A mid-pilot review was completed in January 2023. At this time, there were nine pilot participants. The review involved consulting with pilot stakeholders, reviewing program documentation, and examining pilot data. The results of key performance indicators were included, without any details of the calculation methodology, as well as the numbers of inquiries for the pilot through various communications channels. The report contained quantitative and qualitative analysis to identify the low participation numbers.

42 National Health and Medical Research Council, *Appendix F: Levels of evidence and recommendation grading*, 2009, available from <https://www.nhmrc.gov.au/sites/default/files/images/appendix-f-levels-of-evidence.pdf> [accessed 2 April 2024].

Leigh, A., *What evidence should social policymakers use?*, 2007, p.35 (box 3), available from https://treasury.gov.au/sites/default/files/2019-03/03_Evidence_Hierarchy.rtf [accessed 2 April 2024].

43 A quasi-experimental design where an intervention is assigned to only those above or below an assigned cutoff. Outcomes can be compared above and below the cutoff to determine the causal effect of the intervention. See Australian Centre for Evaluation, *Quasi-experimental methods*, available from <https://evaluation.treasury.gov.au/toolkit/quasi-experimental-methods> [accessed 12 June 2024].

44 A statistical approach where the average change in an outcome over time is compared between a group receiving an intervention and a control group. This approach attempts to mimic an experimental design. See Australian Centre for Evaluation, *Quasi-experimental methods*, available from <https://evaluation.treasury.gov.au/toolkit/quasi-experimental-methods> [accessed 12 June 2024].

3.68 The post-implementation review has not yet been completed (see Table 1.4). There is no methodology that sets out how the post-implementation review will address the research questions identified in the plan, which are as follows:

- Were resources best used?
- Were the outputs satisfactory (e.g. timely support provided to the participants, rehabilitation plans contain sufficient detail, positive impact on participant's wellbeing)?
- Could the outputs have been achieved through an alternative approach?

Department of Home Affairs

3.69 The *Review into Integration, Employment and Settlement Outcomes for Refugees and Humanitarian Entrants in Australia* recommended introducing an employer sponsored visa for skilled refugees (as trialled in the SRLA pilot).⁴⁵ The review stated that for such a program:

- employers should be expected to demonstrate the value of the role being offered;
- employers should propose how they would assist the refugee to integrate into the workforce and community; and
- refugees should be carefully monitored to ensure they are treated consistently with Australia's workplace laws and not economically exploited.

3.70 There was no evidence of Home Affairs applying a robust evaluation methodology that would have provided findings on the outcomes identified above, or any other internally designed outcomes. The minister's brief included four case studies provided by TBB, of pilot participants qualitatively describing their integration into their communities and the value of their role to their employer. There is no evidence of Home Affairs utilising an appropriate methodology to verify or analyse such information and to support the department's recommendation to the minister as to whether to extend the pilot and whether such an extension was a proper use of public resources.

3.71 In November 2023 the Department of Home Affairs advised the ANAO that:

The Department did not consider conducting a formal evaluation process for the Pilot, given the decision for the Minister was whether to extend the Pilot, as opposed to formalising the visa pathway into a permanent arrangement.

Following further discussion and analysis, and consideration of what information (and in what form) would best support the Minister in determining the Pilot's future, the Department formed the view that a less prescriptive and more case- and issue-focussed submission would be a more effective evaluation tool and the 'skeleton' evaluation document was not progressed any further. This was done in consultation with the Minister's Office.

Ethics

3.72 Guidelines for ethical behaviour and decision making in evaluations are specified by the Australasian Evaluation Society. Certain types of research require review by a Human Research Ethics Committee, and when required, research methods must be designed to be consistent with

45 Department of Home Affairs, *Investing in Refugees, Investing in Australia: the findings of a Review into Integration, Employment, and Settlement Outcomes for Refugees and Humanitarian Entrants in Australia*, Home Affairs, Canberra, 2019, available from <https://www.homeaffairs.gov.au/reports-and-pubs/files/review-integration-employment-settlement-outcomes-refugees-humanitarian-entrants.pdf> [accessed 19 March 2024].

ethics approval processes. The key framework for assessing the need for ethics approval is the *National Statement on Ethical Conduct in Human Research (2007)* as set out in the Toolkit.⁴⁶

Department of Health and Aged Care

Naloxone pilot

3.73 Ethics approval was obtained for relevant evaluation activities from the University of Queensland Human Research Ethics Committee (and St Vincent's Hospital Sydney) under the National Mutual Acceptance Scheme. This provided reciprocal approval with public health government Human Research Ethics Committees in New South Wales, South Australia, and Western Australia. Site-specific approvals and additional clearances were obtained where required. All participants provided informed consent prior to engaging in the evaluation process.

Kava pilot

3.74 For the Kava pilot, the need for ethics approval was identified in the evaluation plan and it also stated that the evaluation would be conducted in accordance with the principals of the *National Statement on Ethical Conduct in Research Involving Humans, and Values and Ethics: Guidelines for Ethical Conduct in Aboriginal and Torres Strait Islander Health Research*.⁴⁷

3.75 Multiple ethics approvals were sought at the national and jurisdictional level, as well as to access specific routinely collected data sets.

3.76 One condition of being granted ethical approval from the Australian Institute of Aboriginal and Torres Strait Islander Studies was the need to obtain support from Aboriginal Medical Services Alliance Northern Territory (AMSANT) and Miwatj Health Aboriginal Corporation. AMSANT did not support the involvement of Aboriginal and Torres Strait Islander communities in the NT in this evaluation. The evaluation report indicated that:

AMSANT's concerns related to kava not being identified as a priority in many communities, a lack of local governance in East Arnhem Land, a lack of remuneration [*sic*] for Aboriginal Community Controlled Health Organisations participating in the evaluation, methodology concerns, and potential cultural and social based risks.

3.77 As a result, consultations with members of Aboriginal and Torres Strait Islander communities in the Northern Territory were not able to occur. The implementation of the planned data collection strategy was significantly limited by the lack of ethics approval (see paragraphs 3.13 and 3.56).

Department of Veterans' Affairs

Wellbeing and Support Program (WASP) pilot

3.78 Ethics approval was sought from the Departments of Defence and Veterans' Affairs Human Research Ethics Committee. Ethics approval was granted in accordance with the three-month timeframe estimated in the evaluation plan.

46 National Health and Medical Research Council, *National Statement on Ethical Conduct in Human Research*, NHMRC, Canberra, 2018, available from <https://www.nhmrc.gov.au/research-policy/ethics/national-statement-ethical-conduct-human-research> [accessed 19 March 2024].

47 National Health and Medical Research Council, *Ethical Conduct in research with Aboriginal and Torres Strait Islander Peoples and communities: Guidelines for researchers and stakeholders*, NHMRC, Canberra, 2018, available from <https://www.nhmrc.gov.au/about-us/resources/ethical-conduct-research-aboriginal-and-torres-strait-islander-peoples-and-communities> [accessed 19 March 2024].

Non-Liability Rehabilitation (NLR) program pilot

3.79 Ethics approval is not discussed in the evaluation plan or the mid-pilot review and there was no evidence of consideration of ethical approval for the mid-pilot review.

Department of Home Affairs

3.80 For the SRLA pilot, Home Affairs did not assess whether ethics approval would be required in relation to the evaluation of the pilot.

4. Reporting and advice to the Australian Government

Areas examined

This chapter examines whether there was fit-for-purpose analysis and reporting on evaluation outcomes and whether the advice provided to the Australian Government adequately reflected the learnings from the evaluation.

Conclusion

The Department of Health and Aged Care's (Health) evaluation reporting and advice to the Australian Government for the Take Home Naloxone pilot was largely effective, with the recommendations made to expand the naloxone pilot largely informed by the lessons learnt from the evaluation. Health's evaluation reporting and advice to the Australian Government for the Kava pilot was partly effective as neither the evaluation report nor recommendations on the continuation of the pilot have been provided to the Australian Government.

The evaluation report for the Department of Veterans' Affairs' (DVA) Wellbeing and Support Program pilot was largely effective. There was no evidence of DVA advising the Australian Government on the evaluation findings and impact on future program design. The evaluation for the Non-Liability Rehabilitation pilot has not yet commenced, and reporting and advice to the Australian Government on the mid-pilot review was partly effective.

The Department of Home Affairs' evaluation reporting and advice to the Australian Government for the Skilled Refugee Labour Agreement pilot was partly effective, with outputs rather than pilot outcomes analysed and reported to the minister.

Areas for improvement

The ANAO made one recommendation to the Departments of Veterans' Affairs and Home Affairs aimed at ensuring that advice to government in relation to pilots draws upon evidence and learnings from evaluation.

4.1 Efficient and effective evaluation relies upon not just collecting evidence, but also the accurate analysis and interpretation of the evidence collected, and the subsequent reporting of these findings to responsible decision makers. Evaluations that provide only an account of the process and a description of outcomes will not be sufficient to support informed decision making on pilot extension, ongoing implementation, revision, or closure.

4.2 The Resource Management Guide 130 *Evaluation in the Commonwealth*, also referred to as the Commonwealth Evaluation Toolkit (the Toolkit)⁴⁸ states:

high quality and robust analysis is fundamental to showing how effective the government is, as it provides the evidence to demonstrate that programs and activities are delivering their intended benefits to Australians.

4.3 The Toolkit outlines typical elements of an evaluation report, including that reports: should answer evaluation questions; identify implementation challenges or limitations; and assist decision makers understand whether the program or activity is on track and meeting its objectives. It also

48 Australian Centre for Evaluation, Commonwealth Evaluation Toolkit, available from <https://evaluation.treasury.gov.au/toolkit/commonwealth-evaluation-toolkit> [accessed 1 April 2024].

highlights that the report should make constructive and actionable findings and provide lessons learned to support continuous improvement.

4.4 Appendix 5 contains a summary of the findings from this chapter for each of the pilots examined during the audit.

Was there fit-for-purpose analysis and reporting of pilot evaluation outcomes?

The analysis of pilot evaluation outcomes for Health's pilots and DVA's Wellbeing and Support Program pilot were largely fit for purpose, with the evaluation reports documenting the application of statistical methods to provide defensible findings and making recommendations on the basis of the analysis completed. The evaluation of DVA's Non-Liability Rehabilitation pilot has not yet commenced. Home Affairs' reporting of outputs of the Skilled Refugee Labour Agreement pilot did not contain fit-for-purpose analysis and does not satisfy the requirements of evaluation reporting in the Commonwealth Evaluation Toolkit.

Analysis of pilot outcomes

Department of Health and Aged Care

Naloxone pilot

4.5 Outcomes set out in the evaluation plan for the Naloxone pilot were reported in the first interim report dated 27 April 2020, the second interim report dated 6 November 2020, and the final evaluation report dated 8 September 2021. The first interim period noted amendments to the evaluation plan as a result of early implementation challenges. The second interim report highlighted evaluation challenges for each of the research questions, providing emerging advice from stakeholders to improve implementation.

4.6 The final evaluation report explained the deviations from the evaluation plan and qualified the reporting of outcomes. It also answered all planned research questions and assessed unintended outcomes.

- In terms of pilot implementation, the report confirmed broad alignment with the plan while noting variations across states due to jurisdictional policy structures governing supply arrangements, shortcomings relating to low public awareness, and impacts of unforeseen COVID-19 related import restrictions.
- Data analysis was provided relating to occasions of naloxone supply, registered sites, settings, increase in proportion of people at risk who received naloxone, and percentage of people able to access naloxone within their home postcode.
- The report contained the assessment of the number of overdose reversals per day as a result of access to naloxone supply arrangements under the pilot.
- The impacts on access sites were reported, providing the views of staff, and highlighting staff challenges relating to a lack of funding to assist implementation. Instances of limited engagement with staff in some settings was also highlighted (pain clinics, primary care, and mental health settings).
- Recommendations for a national rollout were clearly described.

4.7 The reporting of evaluation findings was limited by the exclusion of an economic evaluation, which meant no findings could be drawn regarding the economic value of removing barriers to access of naloxone for clients in different settings or jurisdictions. Barriers associated with stigma for opioid users, awareness of naloxone, awareness of overdose risk, fear of police, and negative perceptions of naloxone were not addressed, despite being identified in the evaluation plan as relevant in the context for the pilot. The cost of low levels of awareness and poor uptake in certain settings were not quantified, providing decision makers with no information on the disadvantages of uneven access to naloxone.

Kava pilot

4.8 The analysis detailed in the evaluation report was consistent with that specified in the evaluation plan. The systematic literature review was robust.

4.9 The evaluation was inconclusive around the harms and benefits associated with kava use. The evaluation did not collect information on kava use or imports in the absence of the changes in access introduced as part of the pilot, or how kava use has changed as a result of the pilot. Many of the potential harms relate to long term use, which could not be evaluated over the evaluation period (with data collection occurring between June 2022 and June 2023 after the commencement of phases one and two of the pilot). These risks were not identified in the evaluation plan. The impact of COVID-19 travel bans on personal importation was noted as a limitation on the monitoring data collated and analysed.

4.10 Community surveys, which were the basis of qualitative data on kava use and its harms and benefits, provided information which is not generalisable. Surveys were only undertaken for Pacific Islander communities, and this was not based on random sampling (see paragraphs 3.13 and 3.56). The self-reported data from the community survey of Pacific Islander communities introduced a risk of bias, as most participants viewed the Kava pilot positively, and did not account for this in the analysis. The self-reported impacts were positive, which was counter to the systematic literature review which found that harms and benefits associated with kava consumption varied widely across studies.

4.11 In the absence of direct consultations with identified Aboriginal and Torres Strait Islander communities in the Northern Territory, the scope of the evaluation was adjusted to include additional consultation with government, non-government, and social sector representatives and other agencies. This approach did not provide evidence around uptake or observed harms and was primarily concerned with the pilot process and potential risks of kava use. Several findings of the evaluation were not dependent on the pilot occurring, but rather reflected stakeholder attitudes and concerns around kava in general.

Department of Veterans' Affairs

Wellbeing and support (WASP) pilot

4.12 Qualitative analysis of outcomes of the pilot, and case studies referred to in the evaluation report were largely robust and consistent with the evaluation plan. Case studies included qualitative descriptions and wellbeing measures.

4.13 The quantitative analysis of outcomes in the evaluation was limited due to lower-than-expected data collection throughout the pilot, which was due to risks that were insufficiently or not identified in the evaluation plan (see paragraph 3.16).

- Seventy-three out of 204 veterans who participated in the pilot had both baseline and post pilot outcomes data (35.8 per cent), while 125 (61.3 per cent) had at least baseline data only. There is a risk that the subset of veterans with both baseline and post-pilot outcomes data may have experienced a different change in outcomes compared to pilot participants without baseline and post-pilot outcomes data.
- On average, 34.8 per cent of data was missing due to incomplete questionnaires for baseline variables used to test how the interventions effects on outcomes differ across pilot participants. This approach introduces a risk that the true effect of these baseline variables on the trajectories of the veterans may not be accurate.
- The data on wellbeing and psychological functioning measures did not enable time-series analysis. Linear mixed methods were applied instead (see paragraph 3.63).
- Quality of life indicators were not included in the final evaluation as there was no consistent measure used by health service and rehabilitation providers for this veteran group. The evaluation plan did not state that quality of life measures were not collected in the 12 months before the pilot (see paragraph 3.50) as case managers were concerned about approaching their veteran clients to complete the measures due to their high levels of acuity.
- The healthcare utilisation analysis only used data held by DVA. The evaluation report notes that the dataset used may not have included Medicare Medical Benefits Scheme items or Pharmaceutical Benefits Scheme (PBS) information that were not linked with DVA's system or where the record of health services delivered did not correctly identify the recipient as a veteran.

4.14 Statistical significance was inappropriately assessed in the moderation analysis. The moderation analysis involved testing 56 combinations of baseline variables and wellbeing outcome variables to assess whether there were statistically significant differences in outcome changes depending on characteristics such as marital status. This approach of undertaking many statistical tests and reporting only those that come back with significant results is referred to as data dredging, and significantly elevates the risk of false positives.⁴⁹ There is a risk that this approach led to invalid inferences about statistical significance for the moderation analysis.

4.15 The quantitative analysis was sufficient to draw conclusions about how outcomes had changed relative to baseline for those participating in the pilot. Linking the improvements in outcomes to a causal effect of the pilot was limited by selection bias and the lack of a control group (see paragraph 3.64). The evaluation team could have consulted with those participants with and without baseline and post-pilot outcomes data to estimate the risk of selection bias and the likely impact on evaluation findings. The evaluation report states that linear mixed models used to analyse pre-post changes are robust when data is missing, without establishing the basis for such assumption.

49 The American Statistical Association statement on Statistical Significance and P-values states that 'conducting multiple analyses of the data and reporting only those with certain p-values', 'makes results essentially uninterpretable' and that 'researchers should disclose the number of hypotheses explored during the study, all data collection decisions, all statistical analyses conducted, and all p-values computed' (p.132). See: Wasserstein, R.L. and Lazar, N.A., 'The ASA Statement on p-Values: Context, Process and Purpose', *The American Statistician*, 70(2): 129–133, available from <https://doi.org/10.1080/00031305.2016.1154108> [accessed 2 April 2024].

4.16 The evaluation compared demographic and psychosocial outcomes between WASP pilot participants and the broader veterans and Australian populations. There was no assessment using this data of the extent to which it suggests there would be selection bias among pilot participants compared to other veterans. The evaluation analysis did not involve comparing demographic or other characteristics of pilot participants who only had baseline data recorded to those with data from during or after the pilot. This was a gap in the evaluation that was not noted in the report, with the reader required to assess the reliability of the report findings in light of these limitations.

4.17 The Realist Evaluation Approach focuses on answering the question ‘what works for whom, under what circumstances and how’.⁵⁰ Consistent with this, the evaluation team undertook analysis of how effects of the WASP pilot differ with respect to age, time between discharge and WASP commencement, and time between discharge and the date of the evaluation.

Non-Liability Rehabilitation (NLR) program pilot

4.18 The mid-pilot review states ‘there is no evidence to suggest that the program logic is not valid.’

4.19 The mid-pilot review identified that participation in the NLR pilot was very low and states it is ‘underperforming relative to pre-pilot assumptions’. The mid-pilot review states that the risk of low pilot participation was identified in pilot planning. There was no evidence of this risk being identified and considered as part of evaluation planning.

4.20 The mid-pilot review made 12 recommendations to improve the pilot implementation and evaluation. It provides recommendations for how each gap can be addressed in the remainder of the pilot period. Three recommendations relate specifically to evaluation, generally seeking to incorporate veteran and provider feedback to improve qualitative data, and two recommendations related to establishing benchmarks for client closure data for comparison to other programs.

4.21 In October 2023 DVA advised the ANAO that the post-implementation evaluation of the NLR pilot was due to commence in April 2024.

Department of Home Affairs

4.22 The ministerial brief contained high-level data about program outputs, such as the number of visa applications and grants, and average processing time. In briefing the minister in 2023, the department said:

TBB [Talent Beyond Boundaries]⁵¹ cites high application costs and restrictive legislative criteria, along with limited post arrival settlement support, as primary barriers to the Pilot realising increased uptake.

4.23 Home Affairs did not undertake any data analysis or stakeholder engagement to understand how the experiences from the pilot could inform recommendations or validate the information provided by TBB. The reliance on information from TBB was not cited as a limitation of the analysis contained in the ministerial brief.

50 Public Health England, *A brief introduction to realist evaluation*, 2021, p.3. Available from https://assets.publishing.service.gov.uk/media/60f7fdf7d3bf7f56824cc634/Brief_introduction_to_realist_evaluation.pdf [accessed 1 April 2024].

51 Home Affairs advised the minister that Talent Beyond Boundaries (TBB) was ‘an Australian and international not-for-profit organisation specialising in finding migration pathways for skills migrants and refugees’. TBB supported Home Affairs in delivering the pilot (see paragraph 3.39).

4.24 The Toolkit distinguishes outputs and outcomes, with outputs being ‘tangible deliverables/ products/items delivered by the program or activity, over which management has direct control’. The Toolkit states outcomes can be short-, medium- or long-term, and recommends long-term outcomes be considered as those with ‘wider long-lasting social change’.

4.25 The outcomes of the program include the value of roles filled, integration into communities, and degree of economic exploitation (see paragraph 1.28). The SRLA pilot did not examine the extent to which the pilot achieved these outcomes, deeming the first to be self-evident implicitly for any occupations on the Skilled Occupations List, and not considering the remainder or any other outcomes explicitly.

Fit-for-purpose reporting

Department of Health and Aged Care

Naloxone pilot

4.26 The final report set out a comparison between outcomes identified in the Program Logic versus those achieved, risks that arose and how they were managed, and implications for a national rollout.

4.27 The analysis of the evaluation plan against the recommendations of the Toolkit is set out in Table 4.1.

Table 4.1: Assessment of the Naloxone pilot evaluation report against the Commonwealth Evaluation Toolkit

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Issue or need addressed by the program or activity	◆	The evaluation report sets out the problem of opioid overdose in Australia. The status of naloxone programs in Australia are summarised, and highlights funding challenges, and barriers to access. The pilot's implementation strategy is noted.
Purpose and objectives of the program or activity	◆	
Clear description of how the program is organised and its activities	◆	
Methodology — how the evaluation was conducted and an explanation of why it was done this way. This should include what surveys or interview questions were used and when and how they were delivered (and a copy should be included in the appendix)	▲	The evaluation approach is clearly described, including the Program Logic and Theory of Change that support the methodology. Data sources are well identified, and key evaluation activities are clearly set out. The report did not contain a rationale for why opioid users outside the pilot were not selected as the comparator, nor why an economic evaluation was excluded from scope.

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Sampling — how many people participated in the evaluation, who they were and how they were recruited	◆	The approach to semi-structured interviews for primary data collection was well documented, including the number of individuals recruited, and the number that were interviewed once, twice, or three times. The number of mystery shopping visits was also noted, which were used to observe the settings and practices of naloxone provision.
Data analysis — a description of how data was analysed	◆	The data analysis included comparing access patterns for naloxone supplied through the pilot and access to naloxone through the PBS in pilot and non-pilot jurisdictions. Data analysis of naloxone supply through the pilot examined patterns of access by participant type and setting, which was relevant to the evaluation research questions. Attempts were also made to control for external factors, such as patterns of opioid dispensing (i.e. non-naloxone opioid dispensing).
Ethics — a description of how consent was obtained and how ethical obligations to participants were met	▲	The evaluation report noted that ethics approvals were obtained for all evaluation activities, citing the approvals process, including reciprocal approval. The way that ethical obligations were met was not described.
Findings — what was learnt from the evaluation (and what it means for the program or activity) and how do the results compare with your objective and outcomes	◆	Key findings are accurately summarised in the executive summary.
Findings and recommendations — detailed and actionable suggestions for possible changes to the program or service that have come from the analysis	◆	The implications for a national rollout of the pilot were clearly defined and reflected learnings from the pilot. This included learnings from what was not done well (awareness, especially beyond the alcohol and other drugs sector, training and education), and what is needed for program success (continuity, free access without prescription, broad availability, and coordination across governments).
Any limitations to the evaluation and how future evaluations will overcome these limitations	■	The report documented limitations to the scope of the evaluation. It did not include a 'Research limitations and qualifications' section which outlined the implications of scope limitations. The report did not highlight where the pilot results should be interpreted with care.

Key: ◆ Included in the evaluation report
▲ Partially included in the evaluation report
■ Not included in the evaluation report

Source: ANAO analysis of the Final Evaluation Report.

Kava pilot

4.28 The evaluation report largely contains all elements of evaluation reports recommended by the Toolkit. (see Table 4.2)

Table 4.2: Assessment of the Kava pilot evaluation report against the Commonwealth Evaluation Toolkit

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Issue or need addressed by the program or activity		The introduction to the evaluation report provides an overview of the Kava pilot including the purpose and how the pilot is structured. The purpose of the pilot introduction indicates the likely need for the pilot, or problem being addressed, but does not explicitly describe this in either qualitative or quantitative terms.
Purpose and objectives of the program or activity		
Clear description of how the program is organised and its activities		
Methodology — how the evaluation was conducted and an explanation of why it was done this way. This should include what surveys or interview questions were used and when and how they were delivered (and a copy should be included in the appendix)		The methodology was outlined in the body of the report and detailed further in the appendices, including the number of participants in interviews, focus groups and surveys and how they were recruited. Interview and focus group guides, as well survey questions and information on their mode of delivery are included in the appendix.
Sampling — how many people participated in the evaluation, who they were and how they were recruited		
Data analysis — a description of how data was analysed		
Ethics — a description of how consent was obtained and how ethical obligations to participants were met		Ethics considerations are outlined in the evaluation methodology.
Findings — what was learnt from the evaluation (and what it means for the program or activity) and how do the results compare with the objective and outcomes		Findings are reported against all evaluation questions, although the conclusions for several of them are uncertain.

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Findings and recommendations — detailed and actionable suggestions for possible changes to the program or service that have come from the analysis		The evaluation recommended a monitoring framework which would monitor use and harms over time, without specifically recommending the continuation of the pilot or changes to pilot arrangements.
Any limitations to the evaluation and how future evaluations will overcome these limitations		The report includes a detailed list and explanation of the evaluation's limitations. The report did not state the assumptions used in the economic analysis, as recommended by government guidelines.

Key:  Included in the evaluation report
 Partially included in the evaluation report
 Not included in the evaluation report

Source: ANAO analysis of Health records.

Department of Veterans' Affairs

Wellbeing and support (WASP) pilot

4.29 The evaluation report largely contains all elements of evaluation reports recommended by the Toolkit (see Table 4.3).

4.30 The evaluation report concluded that the weight of evidence from the WASP pilot demonstrated that there was strong approval of the program and that the pilot was successful and improved the lives of veterans.

Table 4.3: Assessment of the WASP pilot evaluation report against the Commonwealth Evaluation Toolkit

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Issue or need addressed by the program or activity		The introduction to the report provides a one-page overview of the issues, needs and key challenges facing veterans and DVA. This includes data about the number and composition of the transitioning veterans cohort and rates of mental health disorders. It summarises the responses of the Australian Government to these challenges and the approach of the WASP pilot.
Purpose and objectives of the program or activity		
Clear description of how the program is organised and its activities		

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Methodology — how the evaluation was conducted and an explanation of why it was done this way. This should include what surveys or interview questions were used and when and how they were delivered (and a copy should be included in the appendix)		The evaluation report has a methods section covering the qualitative analysis methodologies (Realist Evaluation ^a and the RE-AIM framework ^b), the sample and cohort definition, qualitative data collection methods, and quantitative data collection and analysis approaches. The survey tools used in the evaluation are not provided in an appendix.
Sampling — how many people participated in the evaluation, who they were and how they were recruited		The number of veterans participating in the evaluation, their demographic information, a summary of their outcomes at baseline and comparison of these outcomes to other Australian veterans and the general population are provided. The qualitative data collection section states that veterans were recruited via the case managers.
Data analysis — a description of how data was analysed		The report describes quantitative data analysis in detail, providing tables and charts showing counts, averages, standard deviations, and model estimates. The report explains how to interpret results of quantitative data analysis, including discussing certainty of particular conclusions. Description of qualitative data analysis in the report is limited. For example, the differences in views across qualitative data collection methods are not assessed.
Ethics — a description of how consent was obtained and how ethical obligations to participants were met		Administrative details about ethics approval are provided. There is no description about what ethical issues were relevant for the pilot and how ethical obligations were met.
Findings — what was learnt from the evaluation (and what it means for the program or activity) and how do the results compare with your objective and outcomes		Findings from qualitative data analysis are summarised, such as a section describing feedback on the WASP pilot. Findings from quantitative data analysis are drawn throughout the description of quantitative analysis results. A standalone conclusions section draws out findings from the range of analysis elsewhere in the report, including addressing whether the research questions originally posed in the evaluation have been addressed. Findings and conclusions are summarised in the executive summary.

Toolkit's elements of an evaluation report	Included in the evaluation report	ANAO comments
Findings and recommendations — detailed and actionable suggestions for possible changes to the program or service that have come from the analysis		The report made nine recommendations, including that the 'WASP continue in its current, or similar, form'. Recommendations about measurement and future evaluation, such as the recommendation to collect quality of life data using the EQ-5D-5L^c measure, were appropriately based on challenges encountered during the evaluation. The evaluation recommended that further evaluation is needed of the impact on families of veterans with complex care needs, which is appropriate given that the small sample size for this cohort limited the ability of the evaluation to understand these impacts.
Any limitations to the evaluation and how future evaluations will overcome these limitations		The report contains a limitations section, and limitations are summarised in the executive summary. One of the six research questions in the evaluation plan is not included in the questions listed in the report, namely, 'Did health service costs and quality adjusted life years (QALYs) increase as a result of participation with the Pilot?'. The removal of this research question is not noted in the evaluation report. Among the five research questions mentioned in the evaluation report: three were able to be answered; one was not fully determined due to small sample sizes for family members; and one was not able to be determined due to the extremely small sample size of service providers.

Key:  Included in the evaluation report
 Partially included in the evaluation report
 Not included in the evaluation report

Note a: Realistic Evaluation applies a context-mechanism-outcome structure to understand the mechanisms by which the intervention results in changes, and the impact of contextual factors on those mechanisms. See Pawson, R. and Tilley, N. *Realistic Evaluation*, 2024 available from <https://www.urban-response.org/system/files/content/resource/files/main/pawson---tilley-%282004%29-realist-evaluation.pdf> [accessed 1 April 2024].

Note b: RE-AIM determines the population-based impact of a program or policy by assessing interventions across five dimensions: Reach, Effectiveness, Adoption, Implementation and Maintenance. See Glasgow, R.E., Vogt, T.M. and Boles, S.M. *Evaluating the public health impact of health promotion interventions: the RE-AIM framework*, American Journal of Public Health, 1999 September, volume 90, September 1999, pp. 1322–1327. Available from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1508772/> [accessed 1 April 2024].

Note c: The Euroqol (EQ)-5D-5L is a health status measure consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each having five levels.

Source: ANAO analysis of the DVA records.

Non-Liability Rehabilitation (NLR) program pilot

4.31 The NLR pilot evaluation has not yet reached the stage of the post-implementation review and therefore there is no evaluation report to examine.

Department of Home Affairs

4.32 There was no evaluation report for the SRLA pilot. Nevertheless, Home Affairs advised the minister that the aim of the ministerial brief was to ‘provide advice and recommendations to the government on the extension of the Skilled Refugee Pilot and any changes to program settings, governance and monitoring arrangements’.

4.33 The brief provided clear and succinct advice to support the minister’s decision on alterations to the pilot’s processes recommended by Home Affairs. The recommendations were not informed by an evaluation. The absence of data analysis in the ministerial brief reflected the lack of evaluation methodology rather than inadequate reporting of the analysis conducted. An assessment of the brief against the evaluation report recommendations in the Toolkit is set out in Table 4.4.

Table 4.4: Assessment of the advice to the Australian Government on the SRLA pilot outcomes against the Commonwealth Evaluation Toolkit

Toolkit's elements of an evaluation report	Included in ministerial brief titled 'internal evaluation'	ANAO comments
Issue or need addressed by the program or activity		The brief includes background information the issue, purpose and how the pilot is being delivered.
Purpose and objectives of the program or activity		
Clear description of how the program is organised and its activities		
Methodology — how the evaluation was conducted and an explanation of why it was done this way. This should include what surveys or interview questions were used and when and how they were delivered (and a copy should be included in the appendix)		Evaluation questions are not identified, and there is no 'methodology' section.
Sampling — how many people participated in the evaluation, who they were and how they were recruited		No sampling was conducted. To the extent that participants in the pilot are considered to be participants in the internal evaluation, all pilot participants are represented in the data reported, rather than data only being reported for a sample of participants.

Toolkit's elements of an evaluation report	Included in ministerial brief titled 'internal evaluation'	ANAO comments
Data analysis — a description of how data was analysed	■	The number of participants involved, the types of visas granted and the number of days to process the visas are provided. Analysis of such data is limited to calculating averages. The ministerial brief includes data analysis by TBB, including the average salary of pilot participants, retention rate and number of employers by sector. This was prepared by TBB and there no evidence of how Home Affairs gained assurance over the data provided.
Ethics — a description of how consent was obtained and how ethical obligations to participants were met	■	There is no consideration of ethics.
Findings — what was learnt from the evaluation (and what it means for the program or activity) and how do the results compare with your objective and outcomes	■	Findings relate to specific implementation and design challenges. These findings are drawn from general observations of limited performance data rather than from an evaluation exercise. Findings are described along with implications for design choices to be made for the expanded pilot. These findings typically do not derive from data analysis of pilot outcomes, but rather qualitative discussions.
Findings and recommendations — detailed and actionable suggestions for possible changes to the program or service that have come from the analysis	■	Recommendations are made as to continuation of the pilot and changes to its design. The recommendations derive from limited qualitative analysis conducted by TBB rather than from an evaluation activity.
Any limitations to the evaluation and how future evaluations will overcome these limitations	■	No limitations are identified.

Key: ◆ Included in the ministerial brief titled 'internal evaluation'
 ▲ Partially included in the ministerial brief titled 'internal evaluation'
 ■ Not included in the ministerial brief titled 'internal evaluation'

Source: ANAO analysis of Home Affairs records.

Did advice to government and policy proposals draw on pilot evaluation outcomes?

Advice provided by Health to the Australian Government in relation to the Naloxone pilot was appropriate, including the lessons learnt from the pilot. The recommendation to expand the pilot into different environments was partly informed by evaluation. Health has not provided advice to the Australian Government on the findings of the evaluation or lessons learnt in relation to the Kava pilot. DVA did not advise the Minister for Veterans' Affairs on the evaluation findings or lessons learnt for future program delivery for the WASP pilot. Home Affairs' advice to the Australian Government for the continuation of the SRLA pilot was not informed by an evaluation.

Reporting of pilot findings to government

Department of Health and Aged Care

Naloxone pilot

4.34 Health's ministerial submission dated 9 June 2022 referenced the findings of the evaluation, and how it supported scaling up to a national naloxone program. The brief appropriately described the pilot outcomes. It did not refer to limitations of the evaluation or challenges experienced in implementing the pilot. A link to the report published on Health's website was included in an attachment to the brief.

Kava pilot

4.35 The final evaluation report was published on the Office of Drug Control website on 22 December 2023. As at March 2024, no advice had been provided to the Australian Government regarding scale up or adoption of the pilot as a program. As at 19 March 2024, Health's website states 'arrangements under the pilot remain in place until any decision is made by the Australian Government'.

Department of Veterans' Affairs

Wellbeing and Support Program (WASP) pilot

4.36 Evaluation findings were communicated within DVA. DVA has no record of the evaluation report, or any advice on the outcomes of the evaluation or future program design, having been provided to the Australian Government.

4.37 As part of the 2021–22 budget process, DVA sought ongoing funding for the program and its advice to the Australian Government referred to the outcomes set out in the evaluation report. The advice contained a two-paragraph description of the evaluation findings. The advice described the evaluation as finding that the program 'works well', which is an accurate reflection of the report conclusions (see paragraph 4.30).

4.38 The advice to the Australian Government did not identify there were a range of limitations with the evaluation, such as selection bias that may impact on the reliability of the findings.

Non-Liability Rehabilitation (NLR) program pilot

4.39 No advice has been provided to the Australian Government as the post-implementation review has not yet occurred.

Department of Home Affairs

4.40 The ministerial brief containing Home Affairs' recommendations on the continuation of the program was the only reporting to the Australian Government of the outcomes of the pilot. The brief summarised the proposal for changes to the pilot sought by TBB, correspondence from TBB to Home Affairs was also annexed to the brief. The brief did not include a comparison of pilot outcomes relative to what outcomes for those participants would have been under business-as-usual.

4.41 The ministerial brief lists three pilot risks and sensitivities:

- labour agreements can attract public criticism about allowing access by foreign workers to Australian labour markets outside of the usual parameters for the skilled visa program;
- concessions related to candidate documentation give rise to a risk of exploitation by people trying to commit identity or immigration fraud; and
- sponsoring businesses can be subject to integrity risk.

4.42 The brief states that Home Affairs will continue to monitor and minimise any associated risks. No analysis is provided about the size of these risks and how they have materialised during the pilot or would be minimised over the remainder of the pilot and beyond.

Post pilot recommendations and learning from the pilot

Department of Health and Aged Care

Naloxone pilot

4.43 The ministerial submission on the evaluation report appropriately reflected the learnings from the pilot. The brief refers to the funding of \$19.6 million over four years allocated in the 2022–23 budget to deliver a national naloxone program and sought the minister's agreement 'to progress implementation of a national take home naloxone program as announced in the 2022–23 Federal Budget to commence 1 July 2022'. The brief was signed by the minister on 16 June 2022, following the announcement of program funding in the March 2022–23 Federal Budget.

4.44 The recommendation for a national rollout related to access arrangements for naloxone, not to methods for ensuring naloxone was supplied in a cost-effective way. The peak in supply observed during the pilot was traced to a small group of linked pharmacies in one local area that had initially offered naloxone to every patient receiving opioid prescriptions, irrespective of their need or request. This was not addressed in the recommendations in the evaluation report.

4.45 The evaluation demonstrated the pilot's implementation through community and hospital pharmacies was successful. There was limited evidence in the evaluation report to support expansion to non-pharmacy settings, and the minimal involvement of non-pharmacy settings in the pilot was not identified in the ministerial brief. The submission did not highlight elements of the pilot where challenges were experienced, such as poor awareness, especially beyond the alcohol and other drugs sector, inadequate training and education and the impact of the COVID-19 pandemic.

Kava pilot

4.46 In February 2024 Health advised the ANAO that the pilot results have not been presented to the Australian Government, and that it is still considering how pilot findings should be presented in consultation with the Department of Prime Minister and Cabinet.

4.47 There is no evidence that the recommendation of implementing a monitoring framework has been implemented given the continuance of the current pilot arrangements (see paragraph 4.35).

Department of Veterans' Affairs

Wellbeing and support (WASP) pilot

4.48 The pilot was recommended for continuation with limited analysis of the lessons learnt from the evaluation of the pilot (see paragraph 4.38). The evaluation recommended that the EQ-5D-5L measure of quality-of-life be added to the measures collected within the WASP pilot to facilitate future economic evaluation.⁵² DVA's advice to the Australian Government did not mention this recommendation.

Department of Home Affairs

4.49 Home Affairs advised the Australian Government the SRLA pilot should be continued and expanded to 500 places. This was not underpinned by an assessment of pilot outcomes. The brief states:

This will enable the Department to improve and refine policy settings and processes to make the Pilot more attractive to corporate Australia, while managing risk and informing how a larger-scale program might look in the future.

Recommendation no. 6

4.50 The Departments of Veterans' Affairs' and Home Affairs' advice to government on the cessation, continuation or scaling up of a pilot draws on evidence and learnings from the evaluation, including limitations on the robustness of the evaluation undertaken.

Department of Veterans' Affairs' response: Agreed.

4.51 *DVA will further consider opportunities to enhance governance of evaluations for health and wellbeing programs and pilots to facilitate this function, including centralised and consistent records management that captures outcomes of activities and implementation progress. DVA will continue to mature existing evaluation project closure activities to utilise evidence-based evaluation outcomes into decision-making.*

Department of Home Affairs' response: Agreed.

4.52 *The department agrees that all advice on pilot programs with respect to cessation, continuation and scaling up, should draw upon evidence and learnings from an evaluation. This requirement will be communicated to all policy and program areas within the department as part of the broader work to uplift evaluation capability and awareness. In relation to the SRLA Pilot, the department's focus to date has been on fully implementing the Pilot and building a sufficient*

52 The Euroqol (EQ)-5D-5L is a health status measure consisting of five dimensions (mobility, self-care, usual activities, pain and discomfort, and anxiety and depression), each having five levels.

evidence base to assess the viability and future scalability of the program. This process is ongoing. The department will ensure that future advice to government on the cessation, continuation or scaling up the SRLA pilot is supported by an appropriate evaluation.



Rona Mellor PSM
Acting Auditor-General

Canberra ACT
17 June 2024

Appendices

Appendix 1 Entity responses

Department of Health and Aged Care response



Australian Government
Department of Health and Aged Care

Secretary

Ms Rona Mellor PSM
Acting Auditor-General
Australian National Audit Office
GPO Box 707
CANBERRA ACT 2601

Dear  Ms Mellor

Section 19 report on the Australian National Audit Office (ANAO) Evaluation of Australian Government pilot programs

Thank you for your letter of 19 April 2024 to the Department of Health and Aged Care. I appreciate the opportunity to comment on the section 19 report on the *Evaluation of Australian Government pilot programs*.

The department welcomes the ANAO report and recommendations on how our evaluation activity can be improved. The report reflects the commendable work the department has done to date including recent internal frameworks to embed evaluation in investment decision, policy design and program management.

The department continues to improve its evaluation capability through in-house training, enhancement of departmental-specific tools and resources to support high-quality evaluation. Our Strategic Investment Framework embeds evaluation within all our programs to ensure investment decisions are informed by robust and contestable evidence.

The department accepts all recommendations in the audit. Please find the department's summary response at [Attachment A](#) and our responses to each recommendation at [Attachment B](#).

Phone: (02) 6289 8400 Email: Blair.Comley@health.gov.au
Sirius Building, 23 Furzer Street, Woden ACT 2606 - GPO Box 9848 Canberra ACT 2601 - www.health.gov.au

- 2 -

I would like to thank the team from the ANAO for their cooperative and professional manner in working with us in conducting this audit. If you require more information, please contact our Chief Evaluation Officer, Allyson Essex, on 02 6289 9837.

Yours sincerely

A handwritten signature in blue ink, appearing to read 'Blair Comley', with a stylized flourish at the end.

Blair Comley PSM

19 May 2024

Department of Veterans' Affairs response



Australian Government
Department of Veterans' Affairs

SECRETARY

Ref: EC24-000592

Ms Rona Mellor PSM
Acting Auditor-General
Australian National Audit Office
GPO Box 707
Canberra ACT 2601
OfficeoftheAuditorGeneralPerformanceAudit@anao.gov.au

Dear Ms Mellor,

Thank you for your correspondence of 19 April 2024 and the opportunity to respond to the Australian National Audit Office's (ANAO) Proposed Report on the *Evaluation of Australian Government pilot programs* (the Report).

The Department of Veterans' Affairs (DVA) welcomes the audit recommendations and thanks the ANAO for the opportunity to participate. DVA agrees with the four recommendations and acknowledges the opportunities to strengthen frameworks in place to enhance the evaluation of pilots programs.

The department's summary response and the response to recommendations is attached.

My point of contact on this matter is Ms Peta Langeveld, Chief Audit Executive, who can be contacted on 6276 4880, or alternatively Dr Brian Riches, Executive Director, Audit on 6289 6018.

Yours sincerely,

A handwritten signature in black ink, appearing to read 'Alison Frame'.

Alison Frame
Secretary

15 May 2024

Encl. – DVA summary response and response to recommendations

Postal Address: GPO Box 9998 Brisbane QLD 4001 Phone: 1800 VETERAN Web: www.dva.gov.au

Saluting Their Service

Department of Home Affairs response



Australian Government
Department of Home Affairs

SECRETARY

Ms Rona Mellor PSM
A/g Auditor-General
Australian National Audit Office
GPO Box 707
Canberra ACT 2601

Dear Ms Mellor

Thank you for the opportunity to provide feedback and comments on the Section 19 Report for the *Evaluation of pilot programs* that was provided to the Department of Home Affairs on 19 April 2024.

The department agrees with the findings and recommendations and is actively seeking to mature the evaluation culture and practices across the organisation. This includes a focus on embedding evaluation earlier in the policy and program design process, and ensuring decision makers are informed by performance monitoring and evaluation information.

The department recognises the importance of evaluation to support robust evidence in decision making processes and will continue to uplift its evaluation capability and culture as part of broader performance improvement initiatives. Indeed, the evaluation maturity of the department has also been identified as an area for improvement by the Australian Public Service Commission's Capability Review of the department.

The department has work well underway to improve its evaluation capability, including the development of an entity specific enterprise performance framework, the drafting of a decision making criteria to inform evaluation activities and guiding principles for performance monitoring and evaluation activities. The department's central evaluation function is also working closely with the responsible business area to develop an evaluation plan and program in relation to the Skilled Refugee Labour Agreement Pilot (SRLA pilot). The department is engaging with the Commonwealth Evaluation Reference Group, the Australian Centre for Evaluation and counterparts across the Australian Public Service, to inform its work to ensure it is best practice and fit for purpose. This work also includes efforts to improve evaluation maturity as it relates to culture, capability and governance.

The SRLA pilot reviewed in this report commenced in July 2021 as a concept pilot project, designed to demonstrate the viability of complementary labour mobility visa pathways for skilled refugees and displaced people. In line with recommendations made by the ANAO, the department will monitor outcomes and will establish a sufficient evidence base to assess the viability and future scalability of the program. At this stage, no decision has been made on the ongoing future of the SRLA pilot. The department's advice to Government on the future of the SRLA pilot will be informed by an evaluation consistent with the Commonwealth Evaluation Policy.

Please find attached a summary response to the report for inclusion in the formal report (**Attachment A**), as well as a response to the recommendations (**Attachment B**).

Yours sincerely

Stephanie Foster PSM
16 May 2024

Appendix 2 Improvements observed by the ANAO

1. The existence of independent external audit, and the accompanying potential for scrutiny improves performance. Improvements in administrative and management practices usually occur: in anticipation of ANAO audit activity; during an audit engagement; as interim findings are made; and/or after the audit has been completed and formal findings are communicated.
2. The Joint Committee of Public Accounts and Audit (JCPAA) has encouraged the ANAO to consider ways in which the ANAO could capture and describe some of these impacts. The ANAO's Corporate Plan states that the ANAO's annual performance statements will provide a narrative that will consider, amongst other matters, analysis of key improvements made by entities during a performance audit process based on information included in tabled performance audit reports.
3. Performance audits involve close engagement between the ANAO and the audited entity as well as other stakeholders involved in the program or activity being audited. Throughout the audit engagement, the ANAO outlines to the entity the preliminary audit findings, conclusions, and potential audit recommendations. This ensures that final recommendations are appropriately targeted and encourages entities to take early remedial action on any identified matters during the course of an audit. Remedial actions entities may take during the audit include:
 - strengthening governance arrangements;
 - introducing or revising policies, strategies, guidelines, or administrative processes; and
 - initiating reviews or investigations.
4. In this context, the below actions were observed by the ANAO during the course of the audit. It is not clear whether these actions and/or the timing of these actions were planned in response to proposed or actual audit activity. The ANAO has not sought to obtain assurance over the source of these actions or whether they have been appropriately implemented.
 - In September 2023, Home Affairs communicated to all Senior Executive Service officers and Senior Directors a change to its internal approval process for new policy proposals and ministerial and cabinet submissions to support 'aligning their drafting to the Commonwealth Evaluation Policy Statement and the outcomes aligned to the broader APS reform initiatives related to evaluation' (see paragraph 2.30).
 - In January 2024, the Department of Health and Aged Care's (Health) Executive Committee endorsed the creation of a Chief Evaluation Officer position and the introduction of a pre-registration process for Tier one (high-risk) program evaluations. This requires program areas to register the design and analysis plan of an evaluation before it is conducted (see paragraph 2.46).
 - In January 2024, Health entered into a partnership agreement with the Australian Centre for Evaluation, in the Department of the Treasury, with the aim to improve the quality and use of evaluations within Health (see paragraph 2.47).

Appendix 3 Governance arrangements to support evaluation of pilots

1. Table A.1 sets out the summary of the audit's findings in relation to the governance arrangements that support the evaluation of pilots for each entity examined during the audit (see Chapter 2).

Table A.1: ANAO assessment of departmental evaluation arrangements

Evaluation arrangements	Department of Health and Aged Care		Department of Veterans' Affairs		Department of Home Affairs	
	Result	Paragraphs	Result	Paragraphs	Result	Paragraphs
Does the entity have risk-based mechanisms to trigger an evaluation?	◆	2.21 – 2.22	◆	2.23 – 2.28	■	2.29 – 2.31
Do governance arrangements identify appropriate accountabilities for evaluation?	▲	2.38 – 2.40	◆	2.41 – 2.42	■	2.43
Is there a positive culture of evaluation?	◆	2.44 – 2.48	▲	2.49 – 2.51	■	2.53 – 2.54
Are there fit-for-purpose policies and procedures to guide evaluation activities in the entity?	▲	2.59 – 2.60	▲	2.61 – 2.63	■	2.64
Is evaluation specific training available?	▲	2.66 – 2.67	▲	2.68	■	2.69

Key: Yes ◆ Partly ▲ No ■

Source: ANAO analysis.

Appendix 4 Commonwealth Evaluation Principles

1. The Commonwealth evaluation principles are shown in table A.2.

Table A.2 Commonwealth Evaluation Principles

Principle	Description
Fit for purpose	The choice to evaluate and the scale of effort and resources allocated to an evaluation should be proportional to the value, impact, strategic importance and risk profile of the program or activity. Methods should differentiate between evaluations to inform program administration and evaluations to inform policy decisions.
Useful	Evaluations to inform program delivery should be designed for the purposes of continuous improvement and accountability against objectives. Evaluations for decision-making should be designed for the purpose of defining achievable outcomes, taking account of any pilots, prototyping or experience from other jurisdictions. A strong understanding of Government policy intent is required, both when evaluation is used as a monitoring tool and when it is an input to new program design.
Robust, ethical and culturally appropriate	Evaluations should be well-designed, identify potential evaluator bias and take account of the impact of programs and evaluations on stakeholders. Robust data and evidence should provide performance insights and drive continuous improvement for programs in the delivery stage. Ethical and culturally appropriate approaches should be considered in all evaluation activities, including for the collection, assessment, and use of information.
Credible	Evaluations should be conducted by people who are technically and culturally capable. The collection and analysis of evidence should be undertaken in an impartial and systematic way, having regard to the perspectives of all relevant stakeholders. Evaluations should adhere to appropriate standards of integrity and independence.
Transparent where appropriate	To be useful, evaluation findings should be transparent by default unless there are appropriate reasons for not releasing information publicly. To support continuous improvement, accountability and decision-making, evaluation findings should be provided to appropriate stakeholders.

Source: Commonwealth Evaluation Toolkit — 'What is evaluation'.

Appendix 5 Evaluation approach and reporting and advice to the Australian Government

1. Table A.3 sets out the summary of the audit's findings in relation to the evaluation approach, reporting and advice to the Australian Government for the pilots examined during the audit (see Chapters 3 and 4).

Table A.3: ANAO assessment of pilot-specific evaluations

Pilot specific criteria	Department of Health and Aged Care				Department of Veterans' Affairs				Department of Home Affairs	
	Naloxone pilot ^a		Kava pilot ^b		WASP pilot ^c		NLR pilot ^d		SRLA pilot ^e	
	Result	Paragraphs	Result	Paragraphs	Result	Paragraphs	Result	Paragraphs	Result	Paragraphs
Was a fit-for-purpose evaluation plan developed for the pilot?		3.9 – 3.11		3.12 – 3.14		3.15 – 3.17		3.18 – 3.21		3.22 – 3.24
Were stakeholders effectively engaged in relation to the pilot?		3.31		3.32 – 3.34		3.35 – 3.36		3.37 – 3.38		3.39 – 3.41
Was there a review of data sources to support the evaluation design for the pilot?		3.43 – 3.45		3.46 – 3.47		3.48 – 3.50		3.51		3.52
Does the evaluation have a robust methodology that supports informed decision-making?		3.53 – 3.54		3.55 – 3.58		3.59 – 3.65		3.66 – 3.68		3.69 – 3.71
Were ethics appropriately considered in the evaluation method?		3.73		3.74 – 3.77		3.78	N/A	3.79	N/A	3.80

Pilot specific criteria	Department of Health and Aged Care				Department of Veterans' Affairs				Department of Home Affairs	
Was analysis of pilot outcomes robust?	▲	4.5 – 4.7	▲	4.8 – 4.11	▲	4.12 – 4.17	N/A	4.18 – 4.21	■	4.22 – 4.25
Was reporting of the findings fit for purpose?	◆	4.26 – 4.27	◆	4.28	▲	4.29 – 4.30	N/A	4.31	▲	4.32 – 4.33
Were the pilot findings adequately reported to government?	▲	4.34	■	4.35	■	4.36 – 4.38	N/A	4.39	■	4.40 – 4.42
Did post-pilot recommendations to government reflect learnings from the pilot?	◆	4.43 – 4.45	■	4.46 – 4.47	▲	4.48	N/A	–	▲	4.49

Key:

N/A Not applicable

Yes ◆

Partly ▲

No ■

Note a: The Take Home Naloxone pilot (see paragraph 1.9).

Note b: The Kava pilot (see paragraph 1.15).

Note c: The Wellbeing and Support Program pilot (see paragraph 1.22).

Note d: The Non-Liability Rehabilitation pilot (see paragraph 1.25).

Note e: The Skilled Refugee Labour Agreement pilot (see paragraph 1.28).

Source: ANAO analysis.