



## **Supporting Information**

### **Supplementary results**

**This appendix was part of the submitted manuscript and has been peer reviewed.  
It is posted as supplied by the authors.**

Appendix to: Reed RL, Roeger L, Kwok YH, et al. A general practice intervention for people at risk of poor health outcomes: the Flinders QUEST cluster randomised controlled trial and economic evaluation. *Med J Aust* 2022; doi: 10.5694/mja2.51484.

## **1. Participant inclusion and exclusion criteria**

Participants were active patients of the practice (attended the practice on three or more occasions in the previous two years) and assessed (by their regular GP) at risk of poor health outcomes and possibly likely to benefit from the intervention.

Participants were drawn from one of three cohorts:

(1) Older people (aged 65 years and over)

(2) Adults (aged 18 to 64 years) with two or more of the following types of chronic illnesses:

- Cardiovascular (including: ischaemic heart disease, cerebrovascular disease/stroke, peripheral vascular disease, congestive heart failure, hypertension);
- Respiratory (including: asthma, chronic obstructive pulmonary disease);
- Musculoskeletal (including: osteoarthritis, rheumatoid arthritis, other arthritis, osteoporosis, chronic back pain);
- Psychological (including: depression, anxiety, insomnia);
- Digestive (including: gastro-oesophageal reflux disease);
- Kidney disease;
- Diabetes (type 1 or type 2); and
- Cancer

(3) Children and young people (17 years of age or younger).

### **Exclusion criteria:**

Participants were not eligible for the study if they:

- Had a clinically dominant disease likely to shorten life expectancy (e.g. late stage cancer);
- Were unable to comprehend English;
- Suffered from significant cognitive impairment;
- Had a severe mental illness;
- Were not community dwelling (e.g. in a residential aged care facility);
- Were currently enrolled in another health interventional study;
- Were enrolled in supplemental chronic disease management program (e.g. CareFirst, Department of Veterans Affairs Coordinated Veterans' Care Program).

## **2. Operationalisation of the enhanced general practice services**

Practices signed a memorandum of understanding agreeing that, if allocated to the intervention group, they would use their best endeavours to implement the intervention to a high standard. During the study, a research practice nurse (PN) provided a practice facilitation role to help practices with participant recruitment and intervention implementation.

We expected that the multicomponent intervention would be integrated into general practice routine systems of care, with each practice deciding the changes required in their own practices to implement the intervention to a high standard.

To assist the 10 intervention group practices, we arranged a ‘kick-off’ workshop for practice staff. This was well attended by 41 practice staff (15 GPs, 8 PNs, 11 practice managers and 7 administrative officers). The workshop provided a general overview of Flinders QUEST and presented practice-level results from the baseline participant survey (demographics, EQ-5D and health literacy questionnaire). Practice staff brainstormed the best ways to implement the intervention. A second workshop was held for intervention group practices (attended by 29 practice staff: 6 GPs, 7 PNs, 8 practice managers and 8 administrative officers) where the results from the 6 month participant survey were presented. These results were reported at the practice level and included scores on the VAS and HLQ and in addition results from four questions relating to participant’s experiences with respect to wait times for appointments, access to their preferred GP and the length of time that their GP spent with them.

Intervention group practices flagged QUEST participants in their practice software systems. This enabled practice administrative staff to easily identify QUEST participants when they telephoned to make an appointment. QUEST participants were prioritised to receive an appointment with their preferred GP and offered longer appointments. To facilitate this improved access, intervention group practices created appointment times reserved for QUEST participants. Usually, if these had not been filled by midday, the appointment time became available for other patients.

With respect to follow-up after a health event such as an emergency department (ED) presentation or hospitalisation, intervention group PNs or administrative staff checked each day for any discharge summaries received (by fax) against a list of their QUEST participants. If a discharge summary was received for a QUEST participant, the PN would bring this to the attention of the GP and, where clinically indicated, a follow-up face-to-face appointment would be arranged.

Intervention group participants were provided with a laminated QUEST trial card with their name, trial ID, the name of their preferred GP and practice contact details (telephone number). Participants were asked to inform practice staff that they were a QUEST participant when making appointments. They were also asked to inform their practice if they experienced an ED presentation or hospitalisation.

Intervention group practices were asked not to charge co-payments to QUEST participants during the intervention period of the trial. The reason for this was twofold: (1) to remove any disincentive for participants who were charged a co-payment to make appointments with their preferred GP; and (2) to ensure that intervention group participants were not financially disadvantaged through receiving longer appointments and possibly paying higher co-payment charges.

### 3. Protocol changes

1. Change to primary outcome. In the original protocol, the primary outcome for the study was the EQ-5D. When we had recruited approximately 400 participants (prior to randomisation), we changed the primary outcome to the visual analogue scale (VAS) of the EQ-5D. There were two reasons for the change. First, children and young people in the study completed the proxy or youth versions of the EQ-5D questionnaire and the utility scores from these scales are not directly comparable to those from the EQ-5D-5L. Second, taking into account the relatively short time period (12 months) of the intervention, we believed that the VAS would be a more sensitive measure of change.

2. Change to eligibility criteria. Prior to recruitment starting we added an additional eligibility criterion for our participant selection, namely that the patients must be active patients at the practice. Active patients were defined as patients attending the practice on two or more occasions in the previous two years.<sup>1</sup> This could have been assumed from the data extract process we were using, but we believed it should be made explicit.

3. Four questions were added to the 6 and 12-month participant questionnaire. These questions related to participant's experiences with respect to wait times for appointments, access to their preferred GP and the length of time that their GP spent with them. The purpose of these questions was to provide this information to intervention group practices at a 6 month workshop (see part 2 above), and together with qualitative interviews conducted with a subset of participants, gather participant feedback about their general practice experience during the trial.

#### 3. Administrative changes:

- General practice and general practitioner (GP) name added to the expression of interest letter to track response rate to GP invitations to take part in the study.
- Participant contact telephone number added to the expression of interest letter to enable participants to be followed up by telephone after 2 weeks.
- Amendment to allow the use of Research Electronic Data Capture (REDCap) for data collection and storage.
- A question added to the 6-month questionnaire asking whether participants would be agreeable to taking part in a qualitative interview about their experience of the intervention.
- Data-sharing agreement with SA Health expanded to include hospital data from the Central Adelaide Local Health Network.

#### 4. Power calculation

Power was calculated for detecting a standardised effect size on the VAS of between 0.3 and 0.5 standard deviation. Half a standard deviation change is commonly taken to reflect the minimal clinically important difference (MCID) for health-related quality of life<sup>2</sup> but a smaller standardised effect size of 0.3 was thought to have potentially important policy implications at a population level.

On the basis of a South Australian population survey that included older respondents,<sup>3</sup> we assumed a mean (SD) EQ-5D VAS score of 72.7 (17.7) at baseline. We assumed ten general practices (clusters) per arm, a VAS intraclass correlation coefficient of 0.03,<sup>4</sup> a coefficient of variation in cluster size of 0.50 (estimated from previous research<sup>5</sup>) and a VAS pre-post test correlation of 0.6.<sup>6</sup>

With 1000 participants (mean 50 per practice), trial power was estimated (at  $\alpha = 0.05$ , two-sided) using the STATA `clustersampsi` command<sup>7</sup> to be 92% to detect a VAS 0.3-standard deviation change. Within each of the three trial cohorts (assuming an equal distribution of 333 participants), trial power was estimated to be 73% to detect a 0.3-standard deviation change and 99% to detect a 0.5-standard deviation change. A 10% attrition rate was assumed for a total target sample size of 1100.

**Table 1. Characteristics of general practices and general practitioners in the control and intervention groups at baseline**

<b>General practice characteristics</b>	<b>Control</b>	<b>Intervention</b>
No. of practices	10	10
No. of FTE GPs at practice		
1–2	2	1
3–4	2	5
5–6	5	2
7–8	1	1
9–10	0	1
Mean (SD) practice IRSD <sup>A</sup>	5.2 (2.4)	5.7 (3.3)
<b>GP characteristics</b>		
No. of GPs in study	44	48
No. of female GPs	22 (52%)	20 (48%)
No. graduated in Australia	38 (86%)	38 (79%)
Mean (SD) GP years of experience	16.4 (11.5)	18.8 (11.2)
GP hours worked at practice per week (missing = 1)		
11–20	7 (16%)	8 (17%)
21–40	33 (75%)	37 (79%)
41–60	4 (9%)	2 (4%)
GP relation to practice (missing = 3)		
Principal	14 (33%)	21 (45%)
Contractor	17 (41%)	16 (34%)
Employee	7 (17%)	8 (17%)
Registrar	2 (5%)	2 (4%)
Other	2 (5%)	0

<sup>A</sup> Socio-Economic Indexes for Areas, Index of Relative Socio-Economic Disadvantage (IRSD) based on practice suburbs.

FTE = full-time equivalent.

## 5. Sensitivity analysis for the VAS

Sensitivity analyses were performed for the primary outcome (VAS), using alternative specifications for missing data and adjustment for baseline differences.

The missing data analyses used a dataset of completed cases only and a dataset with imputed missing values utilising information for the reason for participant drop-out. The VAS intervention effect (completed cases) in the total cohort was a not statistically significant 0.53-point decrease (coefficient,  $-0.53$ ; 95% CI,  $-2.56$  to  $1.51$ ;  $P = 0.61$ ). The VAS intervention effect (imputed values) in the total cohort was a not statistically significant 0.44-point decrease (coefficient,  $-0.44$ ; 95% CI,  $-2.77$  to  $1.89$ ;  $P = 0.71$ )

Using a cut-point of a small effect size (ES) defined as a standardised difference in means or proportions divided by the standard error greater than 0.10, there were several baseline demographic differences between control and intervention group participants. Compared with the control group, the intervention group had a higher proportion of participants in the older cohort (66.9% v 61.6%, ES = 0.12), a lower proportion of females (50.2% v 57.5%, ES = 0.15), a lower proportion of Aboriginal or Torres Strait Islander participants (0.6% v 2.1%, ES = 0.14), a higher proportion with more than a high school education (52.2% v 45.8%, ES = 0.17), and a higher proportion in a higher ( $> \$40\,000$ ) income category (34.6% v 26.9%, ES = 0.19). The VAS intervention effect adjusted for these covariates (total cohort) was a not statistically significant 0.19-point increase (coefficient,  $-0.19$ ; 95% CI,  $-2.30$  to  $2.34$ ;  $P = 0.99$ ).

**Table 2. Primary and secondary outcomes: children and young people**

	Control		Intervention	
	Baseline	12 months	Baseline	12 months
Primary outcome				
	(n = 28)	(n = 27)	(n = 30)	(n = 27)
VAS score, mean (SD) <sup>A</sup>	78.3 (20.9)	80.6 (20.7)	79.4 (18.5)	81.4 (15.8)
Secondary outcomes				
Hospital service use <sup>B</sup>	(n = 27)	(n = 27)	(n = 30)	(n = 30)
ED presentations, mean (SD)	1.30 (2.76)	0.93(1.59)	1.10 (1.81)	0.93 (1.70)
Admissions, mean (SD)	0.59 (1.31)	0.37 (0.74)	0.53 (1.07)	0.43 (1.19)
Total stay (nights), mean (SD)	3.00 (1.66)	0.44 (1.45)	0.90 (2.96)	1.23 (5.13)
	(n = 25)	(n = 25)	(n = 28)	(n = 28)
Medicare specialist claims, mean (SD) <sup>C</sup>	13.0 (23.6)	12.7 (14.1)	10.2 (10.8)	10.3 (14.7)
PBS items supplied, mean (SD) <sup>C</sup>	3.56 (3.61)	6.08 (4.09)	7.93 (7.54)	8.25 (9.69)

Data are given as summary statistics of the mean and standard deviation (SD) excluding missing data.

<sup>A</sup> Scores on the VAS range between 0 ‘the worst health you can imagine’ to 100 ‘the best health you can imagine’. The VAS dataset comprises the child cohort of 58 participants.

<sup>B</sup> The hospitalisation dataset comprises 57 participants (98.3% of the 58 participants in the child cohort) matched to the SA Health Patient Master Index.

<sup>C</sup> The Medicare/PBS dataset comprises 53 participants (91.4% of the 58 participants in the child cohort) matched to Services Australia records. The analysis dataset excludes GP items.

For the VAS, ‘baseline’ is at the start of the intervention. For hospital utilisation, Medicare claims and PBS items supplied, ‘baseline’ is for the 12-month period preceding the intervention. An intervention effect is not estimated because the sample size was insufficient. The HLQ was not completed by the child cohort. Cost-effectiveness was not calculated for the child cohort.

ED = emergency department; PBS = Pharmaceutical Benefits Scheme; VAS = Visual Analogue Scale of the EQ-5D-3L.

**Table 3. Primary and secondary outcomes: adult cohort**

	Control		Intervention		Intervention effect	
	Baseline	12 months	Baseline	12 months		<i>P</i>
Primary outcome	( <i>n</i> = 171)	( <i>n</i> = 150)	( <i>n</i> = 144)	( <i>n</i> = 132)	Coefficient (95% CI)	
VAS score, mean (SD) <sup>A</sup>	62.5 (21.5)	66.2 (22.0)	62.7 (20.2)	65.1 (20.8)	-1.22 (-5.05–2.61)	0.53
Secondary outcomes						
Hospital service use <sup>B</sup>	( <i>n</i> = 169)	( <i>n</i> = 169)	( <i>n</i> = 141)	( <i>n</i> = 141)	IRR (95% CI)	
ED presentations, mean (SD)	0.73 (1.51)	0.68 (1.68)	0.78 (1.95)	0.77 (2.29)	1.05 (0.63–1.76)	0.85
Admissions, mean (SD)	0.49 (1.13)	0.46 (1.27)	0.37 (0.81)	0.43 (1.12)	1.25 (0.68–2.31)	0.48
Total stay (nights), mean (SD)	1.64 (9.03)	1.69 (7.47)	1.00 (3.52)	2.04 (14.78)	0.82 (0.20–3.38)	0.79
Medicare specialist claims, mean (SD) <sup>C</sup>	( <i>n</i> = 169)	( <i>n</i> = 169)	( <i>n</i> = 143)	( <i>n</i> = 143)	IRR (95% CI)	
PBS items supplied, mean (SD) <sup>C</sup>	30.1 (28.2)	25.3 (19.8)	33.7 (31.9)	29.9 (27.9)	0.99 (0.84–1.16)	0.88
Health Literacy Questionnaire <sup>D</sup> , mean score (SD)	48.2 (37.1)	48.41 (38.3)	52.50 (38.6)	51.99 (38.2)	1.00 (0.92–1.09)	0.94
Feeling understood and supported by healthcare providers	( <i>n</i> = 170)	( <i>n</i> = 150)	( <i>n</i> = 144)	( <i>n</i> = 132)	Coefficient (95% CI)	
Feeling understood and supported by healthcare providers	3.43 (0.49)	3.38 (0.61)	3.41 (0.56)	3.44 (0.52)	0.08 (-0.05–0.20)	0.23
Have sufficient information to manage health	3.06 (0.56)	3.12 (0.63)	2.97 (0.65)	3.13(0.58)	0.07 (-0.06–0.21)	0.30
Ability to actively engage with healthcare providers	3.93 (0.71)	3.95 (0.69)	3.90 (0.66)	4.01 (0.71)	0.08 (-0.07–0.22)	0.31
Navigating the health care system	3.73 (0.68)	3.78 (0.77)	3.60 (0.64)	3.82 (0.68)	0.16 (0.01–0.30)	0.04
Understanding information well enough to know what to do	4.06 (0.69)	4.16 (0.61)	3.98 (0.67)	4.16 (0.58)	0.06 (-0.06–0.19)	0.31

Unless indicated otherwise, data are given as summary statistics of the mean and standard deviation (SD) excluding missing data.

<sup>A</sup> Scores on the VAS range between 0 ‘the worst health you can imagine’ to 100 ‘the best health you can imagine’. The VAS dataset comprises the adult cohort of 315 participants. Intervention effect is estimated from a multilevel linear regression model.

<sup>B</sup> The hospitalisation dataset comprises 310 participants (98.4% of the 315 participants in the adult cohort) matched to the SA Health Patient Master Index. Intervention effect is estimated from a multilevel negative binomial model. Total night stays was a post-hoc analysis.

<sup>C</sup> The Medicare/PBS dataset comprises 312 participants (99.0% of the 315 participants in the adult cohort) matched to Services Australia records. Intervention effect is estimated from a multilevel negative binomial model. The analysis dataset excludes GP items.

<sup>D</sup> Higher Health Literacy Questionnaire scores indicate better health literacy. Scales: Feeling understood and supported by healthcare providers (4 items, range 1–4); Have sufficient information to manage health (5 items, range 1–5); Ability to actively engage with healthcare providers (5 items, range 1–5); Navigating the health care system (6 items, range 1–5); Understanding information well enough to know what to do (5 items, range 1–5). The Health Literacy Questionnaire dataset comprises the adult cohort of 315 participants minus 1 with missing data. Intervention effect is estimated from a multilevel linear regression model.

For the VAS, HLQ and EQ-5D-5L, ‘baseline’ is at the start of the intervention. For hospital utilisation, Medicare claims and Pharmaceutical Benefits Scheme items supplied, ‘baseline’ is for the 12-month period preceding the intervention.

CI = confidence interval; ED = emergency department; ICER = incremental cost-effectiveness ratio; IRR = incidence rate ratio; SD = standard deviation; VAS = Visual Analogue Scale of the EQ-5D-5L.

**Table 4. Primary and secondary outcomes: older adult**

	Control		Intervention		Intervention effect	
	Baseline	12 months	Baseline	12 months		<i>P</i>
Primary outcome	( <i>n</i> = 319)	( <i>n</i> = 296)	( <i>n</i> = 352)	( <i>n</i> = 332)	Coefficient (95% CI)	
VAS score, mean (SD) <sup>A</sup>	72.2 (18.2)	72.3 (18.5)	74.0 (17.9)	74.9 (18.0)	0.43 (−2.02–2.87)	0.73
Secondary outcomes						
Hospital service use <sup>B</sup>	( <i>n</i> = 313)	( <i>n</i> = 313)	( <i>n</i> = 348)	( <i>n</i> = 348)	IRR (95% CI)	
ED presentations, mean (SD)	0.58 (1.21)	0.71 (1.50)	0.64 (1.62)	0.59 (1.27)	0.80 (0.58–1.12)	0.20
Admissions, mean (SD)	0.45 (0.96)	0.61 (1.21)	0.46 (1.07)	0.48 (1.06)	0.76 (0.52–1.11)	0.16
Total stay (nights), mean (SD)	1.66 (6.07)	1.96 (5.26)	1.42 (5.58)	1.28 (4.31)	0.55 (0.26–1.15)	0.11
Medicare specialist claims, mean (SD) <sup>C</sup>	( <i>n</i> = 307)	( <i>n</i> = 307)	( <i>n</i> = 343)	( <i>n</i> = 343)	IRR (95% CI)	
	41.7 (41.1)	42.8 (59.5)	40.2 (31.8)	40.2 (38.4)	1.00 (0.90–1.12)	0.97
PBS items supplied, mean (SD) <sup>C</sup>	60.5 (35.6)	60.1 (35.0)	57.5 (35.6)	57.2 (35.3)	1.00 (0.97–1.04)	0.97
Health Literacy Questionnaire <sup>D</sup> , mean score (SD)	( <i>n</i> = 315)	( <i>n</i> = 295)	( <i>n</i> = 345)	( <i>n</i> = 332)	Coefficient (95% CI)	
Feeling understood and supported by healthcare providers	3.48 (0.51)	3.48 (0.48)	3.51 (0.51)	3.55 (0.50)	0.03 (−0.05–0.11)	0.47
Have sufficient information to manage health	3.21 (0.57)	3.26 (0.54)	3.24 (0.53)	3.28 (0.53)	−0.02 (−0.11–0.06)	0.59
Ability to actively engage with healthcare providers	4.25 (0.57)	4.26 (0.57)	4.28 (0.59)	4.29 (0.58)	−0.01 (−0.09–0.07)	0.73
Navigating the health care system	4.04 (0.58)	4.08 (0.57)	4.10 (0.61)	4.12 (0.55)	−0.02 (−0.10–0.06)	0.66
Understanding information well enough to know what to do	4.17 (0.65)	4.20 (0.66)	4.16 (0.62)	4.19 (0.61)	0.00 (−0.08–0.08)	0.98

Unless indicated otherwise, data are given as summary statistics of the mean and standard deviation (SD) excluding missing data.

<sup>A</sup> Scores on the VAS range between 0 ‘the worst health you can imagine’ to 100 ‘the best health you can imagine’. The VAS dataset comprises the older cohort of 671 participants. Intervention effect is estimated from a multilevel linear regression model.

<sup>B</sup> The hospitalisation dataset comprises 661 participants (98.5% of the 671 participants in the older cohort) matched to the SA Health Patient Master Index. Intervention effect is estimated from a multilevel negative binomial model. Total night stays was a post-hoc analysis.

<sup>C</sup> The Medicare/PBS dataset comprises 650 participants (96.9% of the 671 participants in the older cohort) matched to Services Australia records. Intervention effect is estimated from a multilevel negative binomial model. The analysis dataset excludes GP items.

<sup>D</sup> Higher HLQ scores indicate better health literacy. Scales: Feeling understood and supported by healthcare providers (4 items, range 1–4); Have sufficient information to manage health (5 items, range 1–5); Ability to actively engage with healthcare providers (5 items, range 1–5); Navigating the health care system (6 items, range 1–5); Understanding information well enough to know what to do (5 items, range 1–5). The HLQ dataset comprises the older cohort of 671 participants minus 11 with missing data. Intervention effect is estimated from a multilevel linear regression model.

For the VAS, HLQ and EQ-5D-5L, ‘baseline’ is at the start of the intervention. For hospital utilisation, Medicare claims and PBS items supplied, ‘baseline’ is for the 12-month period preceding the intervention. CI = confidence interval; ED = emergency department; HLQ = Health Literacy Questionnaire; ICER = incremental cost-effectiveness ratio; IRR = incidence rate ratio; PBS = Pharmaceutical Benefits Scheme; SD = standard deviation; VAS = Visual Analogue Scale of the EQ-5D-5L.

**Table 5. Medicare specialist claims, by category**

	Control		Intervention		Intervention effect <sup>A</sup>	
	Baseline	12 months	Baseline	12 months	IRR (95% CI)	P
No. of participants	501		514			
Category	36.3 (37.2)	35.4 (49.0)	36.8 (31.8)	35.7 (35.6)	1.00 (0.91–1.09)	0.94
Diagnostic	1.60 (2.28)	1.54 (2.20)	1.63 (2.22)	1.40 (1.87)	0.91 (0.76–1.08)	0.29
Therapeutic	1.85 (8.50)	1.64 (8.65)	2.12 (7.95)	1.51 (4.94)	0.74 (0.46–1.21)	0.23
Surgical	1.39 (2.60)	1.26 (2.42)	1.69 (2.78)	1.53 (3.23)	0.94 (0.74–1.20)	0.64
Imaging	3.47 (4.22)	3.13 (4.32)	3.17 (4.04)	3.10 (3.66)	1.11 (0.93–1.32)	0.25
Pathology	19.4 (22.6)	19.2 (28.3)	19.8 (18.8)	19.6 (20.6)	1.04 (0.93–1.16)	0.48
Optometry <sup>B</sup>	0.83 (0.96)	0.82 (0.85)	0.85 (0.96)	0.89 (0.96)	1.06 (0.88–1.29)	0.52
Dental	0.12 (0.97)	0.14 (1.07)	0.08 (0.81)	0.04 (0.37)	–	–
Psychology/psychiatry	0.89 (3.27)	0.82 (2.65)	0.78 (2.88)	0.80 (4.04)	0.85 (0.55–1.31)	0.47
Specialists	4.32 (6.58)	4.42 (9.97)	4.25 (5.49)	4.41 (8.53)	0.98 (0.84–1.14)	0.79
Allied health	2.46 (2.79)	2.45 (2.80)	2.47 (3.03)	2.42 (2.81)	1.00 (0.86–1.16)	0.98

Data are the mean (SD) of the number of claims.

<sup>A</sup>The intervention effect (incidence rate ratio [IRR]) is calculated from a multilevel negative binomial regression model.

<sup>B</sup>The optometry intervention effect is calculated from a multilevel Poisson regression model.

Where the intervention effect is not reported, this is due to insufficient sample size. The dataset comprises  $n = 1015$  (97.2%) of the 1044 participants, who could be matched to Services Australia records.

Baseline, the 12-month period prior to the intervention; 12 months, the 12-month intervention period.

CI = confidence interval

**Table 6. Medicare general practitioner (GP) claims, by type**

	Control		Intervention		Intervention effect <sup>A</sup>	
	Baseline	12 months	Baseline	12 months	IRR (95% CI)	P
No. of participants	501		514			
GP claim type (total)	18.13 (11.1)	17.34 (11.3)	17.48 (10.0)	17.92 (10.5)	1.07 (1.02–1.14)	0.012
Brief consult	0.51 (2.11)	0.60 (2.51)	0.59 (1.47)	0.65 (2.05)	0.84 (0.63–1.11)	0.22
Standard consult	9.00 (6.60)	8.06 (6.81)	8.16 (5.41)	7.00 (5.39)	0.96 (0.89–1.03)	0.27
Long consult	3.04 (3.70)	2.96 (3.58)	3.28 (3.93)	3.78 (4.07)	1.21 (1.08–1.36)	0.001
Prolonged consult	0.41 (1.30)	0.36 (1.14)	0.42 (1.64)	0.35 (1.26)	0.96 (0.66–1.38)	0.81
RACF consult	0.09 (1.27)	0.28 (2.39)	0.00 (0.04)	0.14 (1.31)	–	–
GP mental health	0.42 (1.53)	0.30 (1.07)	0.25 (0.86)	0.23 (0.96)	1.23 (0.83–1.81)	0.30
GP acupuncture	0.11 (1.40)	0.14 (1.66)	0.03 (0.27)	0.12 (1.32)	–	–
Home/institution	0.38 (2.27)	0.36 (2.19)	0.10 (0.62)	0.17(0.96)	1.92 (0.78–4.73)	0.16
AH consult	0.29 (0.79)	0.24 (0.70)	0.39 (1.11)	0.39 (1.29)	1.17 (0.82–1.67)	0.39
GP PIP	0.16 (0.37)	0.16 (0.38)	0.25 (0.45)	0.21 (0.42)	0.72 (0.38–1.36)	0.31
GP chronic disease	3.71 (3.27)	3.89 (3.31)	4.01 (3.54)	4.89 (3.93)	1.17 (1.06–1.28)	0.001

Data are the mean (SD) of the number of claims.

<sup>A</sup> The intervention effect (incidence rate ratio [IRR]) is calculated from a multilevel negative binomial regression model.

<sup>B</sup> The optometry intervention effect is calculated from a multilevel Poisson regression model.

Where the intervention effect is not reported, this is due to insufficient sample size. The dataset comprises  $n = 1015$  (97.2%) of the 1044 participants, who could be matched to Services Australia records. GP item classifications are shown in Supporting Information, table 7.

Baseline, the 12-month period prior to the intervention; 12 months, the 12-month intervention period.

AH = after hours; CI = confidence interval; Consult = consultation; PIP = Practice Incentives Program; RACF = residential aged care facility.

**Table 7. Classification of Medicare general practitioner (GP) claims**

<b>GP item classification</b>	<b>Medicare item number</b>
Brief consult	3, 52
Standard consult	23, 53
Long consult	36, 54
Prolonged consult	44, 57
RACF consult	35, 43, 51, 5010, 5028, 5049, 5067, 90020, 90035, 90043, 90051, 90093
GP mental health care	282, 2700, 2701, 2712, 2713, 2715, 2717, 2721, 2725,
GP acupuncture	193, 197, 225, 229, 230, 233, 763, 776
Home/institution consult	4, 24, 37, 47, 59, 60, 65, 507, 511, 5023, 5043, 5063, 5223, 5227
After-hours consult	585, 588, 591, 594, 599, 600, 5000, 5020, 5203, 5207, 5040, 5060, 5208
GP PIP	2501, 2504, 2517, 2521, 2525, 2526, 2546, 2547, 2552, 2558, 2603, 2620, 2664
Chronic disease items	
GP chronic – review	732
GP medication review	245, 900, 903
GP case conferences	699, 735, 747, 750
GP practice nurse	10983, 10997
GP chronic – GPMP	721
GP chronic – TCA	723, 731
GP health assessments	701, 703, 705, 707, 715,

Medicare GP claims were classified based on the Bettering the Evaluation of Care and Health program<sup>5</sup> adapted to accommodate several new items and with an additional combination category of ‘Chronic disease items’ created post-hoc.

GPMP = GP management plan; PIP = Practice Incentives Program; RACF = residential aged care facility; TCA = team care arrangement.

## 6. Economic evaluation

The economic evaluation compared the relative cost-effectiveness of the intervention to usual care (the control). Best-practice guidelines based on the Consolidated Health Economic Evaluation Reporting Standards<sup>8</sup> were followed. The analysis was conducted from an Australian public health provider perspective and included only costs borne by the Australian Government and the South Australian health department.

The analysis was based on a 1-year time horizon, with costs and outcomes between the two groups compared from the time of randomisation to the last follow-up (12 months). Discounting of costs and effectiveness measures was not performed because the time horizon of this study did not exceed 1 year.<sup>9</sup>

### *Choice of outcomes*

The primary outcome for the economic evaluation was the number of quality-adjusted life-years (QALYs) gained over 12 months, calculated using the trapezium method<sup>10</sup> based on responses to the EQ-5D-5L questionnaire at baseline and 12 months. The EQ-5D-5L is a generic health-related quality-of-life measure for use on individuals aged  $\geq 18$  years and consists of five single-item dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression.<sup>11</sup> Its validity, when used in differentiated populations, has been demonstrated.<sup>12-14</sup>

Derived from the original three-level EQ-5D (EQ-5D-3L) questionnaire, the EQ-5D-5L consists of five rather than three levels of impairment in each domain (i.e. no, slight, moderate, severe and extreme problems in the relevant dimension of health). Up to 3125 different health states can be described using responses to the instrument. Utility values, ranging from  $-0.676$  to  $1$ , were estimated using an Australian value set.<sup>15</sup> A utility score of  $1$  represents a 'full health' state, whereas a utility score  $< 0$  represents health states that are worse than death.<sup>16</sup>

Participants in the child cohort completed the proxy or youth versions of the EQ-5D questionnaire. Utility scores from these scales are not directly comparable to those from the EQ-5D-5L<sup>17</sup> and, for this reason, the child cohort was not included in the economic analysis.

### *Estimating resources and cost*

Resource use and costs were estimated from the Australian public health provider perspective. Out-of-hospital Medicare resource use and costs were estimated from Medicare data (to calculate costs of primary care visits, medical consultations, treatments, investigations, and allied health visits). The cost of pharmaceuticals was estimated using PBS data. Costs of ED presentations and admissions to public hospitals were estimated using Australian refined diagnosis-related groups (AR-DRG) data.

The cost of the intervention was set at \$1000 per participant. This was the quantum of payment received by practices for providing the multicomponent intervention to study participants. All costs are reported in Australian dollars at 2019–20-unit prices.

### *Analytical methods*

Analyses were conducted in Microsoft Excel and Stata version 16.1.

### *Base case analysis*

The primary outcome was calculated using responses to the EQ-5D-5L questionnaire. Best-practice economic evaluation was conducted to establish whether the intervention was value for money compared against usual care. An incremental approach was used to determine, where appropriate, incremental cost-effectiveness ratios (ICERs), expressed as the incremental cost per QALY gained. ICERs were calculated by dividing the difference in total costs (incremental costs) by the difference in the QALY gains.<sup>10</sup> An intention-to-treat approach<sup>8</sup> was taken in the analysis.

The statistical analysis was based on a linear mixed model (LMM) estimated using the maximum likelihood estimation approach and conducted under the assumption that any missing data were missing at random.<sup>18</sup> Inferences based on the LMM are valid when data are assumed to be missing at random. Analyses comparing QALY gains over the trial period also controlled for baseline differences in EQ-5D-5L utility within the LMM by including the baseline EQ-5D-5L score as a covariate.<sup>19</sup>

Within-trial economic evaluation with respect to QALYs was conducted, allowing for bivariate uncertainty with bootstrapping of participant costs and effects to maintain the covariance structure. Mean (SEM) and mean differences in costs and outcomes are reported with 95% bootstrapped confidence intervals. Non-parametric bootstrapping<sup>20</sup> was used to determine 1000 paired estimates of mean differences in costs and outcomes from participant-level data. These bootstrapped pairs are presented as cost-effectiveness planes.<sup>21</sup>

To further characterise the uncertainty in the economic evaluation results, cost-effectiveness acceptability curves were constructed. These cost-effectiveness acceptability curves show the probability of the intervention arm being cost-effective compared with the control arm at different willingness-to-pay (WTP) thresholds. A WTP threshold of \$50 000 per QALY gained was used. This is the implicit criterion used for assessing the cost-effectiveness of new pharmaceuticals and medical services in Australia.<sup>22</sup>

#### *Subgroup analysis*

Subgroup analyses were performed for the adults (participants aged 18–64 years with two or more chronic illnesses) and older participants aged  $\geq 65$  years.

#### *Threshold and sensitivity analyses*

Two threshold analyses were conducted. The first determined how low or high the cost of the intervention needed to be in the 1-year post-baseline period for it to be cost-effective, at the threshold of \$50 000 per QALY gained, if its effectiveness remained unchanged. The second examined longer-term costs and outcomes by assessing the impact of varying the time horizon over which the effectiveness of the intervention could be maintained (2- to 10-years post-baseline) as well as the level of effectiveness (i.e. 100%, 75% and 50% of the original efficacy). In this second analysis, the cost of the intervention was only included in the intervention year. For post-intervention periods, costs comprised just those for SA Health hospital, Medicare and PBS resource use. Post-baseline 2-year costs and QALYs were discounted at 5% as per Australian recommendations.<sup>23</sup>

In the sensitivity analysis, the effect of including additional study costs on the cost-effectiveness results was explored. These additional costs were \$10 000 per practice payment in both the intervention and control groups (to cover administrative costs associated with patient recruitment, data collection activities and taking part in qualitative interviews at the conclusion of the trial) and \$4500 per practice in the intervention group only (for a research nurse to provide practice facilitation estimated at 0.5 full-time equivalent [FTE]).

#### *Descriptive statistics and missing data*

A total of 986 participants (490 in the intervention group and 496 in the control group) formed the initial analysis dataset. Missing data across the data sources is summarised in Supporting Information, table 8.

**Table 8. Missing data in cost-effectiveness data sources (adults and older adult groups)**

Data source	Control		Intervention	
	Baseline	12 months	Baseline	12 months
No. of participants		490		496
EQ-5D-5L				
Complete	490 (100)	446 (91.0)	496 (100)	464 (93.5)
Death (utility set to zero)	0 (0.0)	10 (2.0)	0 (0.0)	12 (2.4)
Missing value on item	0 (0.0)	1 (0.2)	0 (0.0)	0 (0.0)
Lost to follow-up	0 (0.0)	33 (6.7)	0 (0.0)	20 (4.0)
Medicare/PBS				
Matched	476 (97.1)	476 (97.1)	486 (98.0)	486 (98.0)
Not matched	14 (2.9)	14 (2.9)	10 (2.0)	10 (2.0)
SA Health hospital records				
Matched	482 (98.4)	482 (98.4)	489 (98.6)	489 (98.6)
Not matched	8 (1.6)	8 (1.6)	7 (1.4)	7 (1.4)
No. of ED presentations	306	336	331	315
No. with cost data (%)	297 (97.1)	332 (98.8)	318 (96.1)	301 (95.6)
No. of hospital admissions	223	267	212	228
No. with cost data (%)	223 (100)	266 (99.6)	209 (98.6)	222 (97.4)

Unless indicated otherwise, data are given as *n* (%).

For the EQ-5D-5L, 'baseline' is at the start of the intervention. For Medicare/PBS and SA Health hospital records, 'baseline' is for the 12-month period preceding the intervention.

ED = emergency department; PBS = Pharmaceutical Benefits Scheme.

For the EQ-5D-5L, there was complete baseline data. At the 12-month follow-up, EQ-5D-5L data were obtained from 910 participants. The reason for missing EQ-5D-5L data in 22 cases was 'Death', and the utility score for these cases was set to zero. Therefore, the dataset for the outcomes component of the cost-effectiveness analysis comprised 932 participants, or 94.5% of the total number of 986 participants in the adults and older cohorts. Missing EQ-5D-5L data rates in the control and intervention groups were 6.9% and 4.0%, respectively.

For the Medicare and PBS data, some missing data occurred through participants not being able to be matched to Medicare records. However, rates of matching were high, with little difference in matching rates between the control (97.1%) and intervention (98.0%) groups.

For the SA Health public hospital data source, 971 participants (98.5%) in the adults and older cohorts could be matched to SA Health records. The match rates were similar between the control (98.4%) and intervention (98.6%) groups. There were 2218 hospital events (ED presentations and hospital admissions) for the 971 matched adult and older cohort participants at baseline (the 12-month period prior to the intervention) and the 12-month period of the intervention. Of these 2218 hospital events, 2168 (97.7%) had cost data. The primary reason for missing cost data was that the hospital was not part of the Southern Adelaide or Central Adelaide Local Health Network, where data sharing agreements enabled cost information to be readily extracted.

Across the Medicare, PBS and SA Health data sources, cost data were missing for 67 participants (6.8%). Therefore, the dataset for the cost component of the cost-effectiveness analysis comprised 919

participants or 93% of the total number of 986 participants in the adults and older cohorts. Missing cost data rates were similar between the control (6.3%) and intervention (7.3%) groups.

### *Results*

At the 12-month follow-up, mean total costs per participant in the adults and older cohorts were higher in the intervention than in the control group (by \$2201 per participant). This difference was not statistically significant (95% CI, -\$1765 to \$6166;  $P = 0.277$ ). The major cost drivers of the difference were the intervention cost (\$1000 per participant) and hospital admission costs (which were \$728 per participant higher in the intervention group; 95% CI, -\$3641 to \$5095;  $P = 0.744$ ) (table 9).

The intervention was more effective than usual care in terms of EQ-5D-5L-based QALYs because it was associated with 0.032 more QALYs gained per participant, and this difference was statistically significant (95% CI, 0.001 to 0.063;  $P = 0.042$ ). The resulting ICER was estimated at \$69 585 per QALY gained (95% CI, \$22 968 to \$116 201), which is above the \$50 000 per QALYs gained threshold for determining cost-effectiveness in Australia. Therefore, in the combined adults and older cohort, the intervention was not cost-effective (table 9).

### *Subgroup analysis*

In the adults cohort, the intervention was more expensive than the control (by \$5559 per participant), but this cost difference was not statistically significant (95% CI, -\$6166 to \$17 283;  $P = 0.353$ ). The intervention was more effective than usual care in terms of EQ-5D-5L-based QALYs because it was associated with 0.007 more QALYs gained per participant, but this difference was not statistically significant (95% CI, -0.068 to 0.082;  $P = 0.847$ ). The resulting ICER was estimated at \$841 202 per QALY gained (95% CI, -\$170 447 to \$1 852 850), which is higher than the \$50 000 per QALY gained threshold for determining cost-effectiveness in Australia. Therefore, in the adults cohort, the intervention was not cost-effective (table 9).

For the older cohort, the intervention was also more expensive than usual care but only by \$571 per participant. The intervention was more effective than usual care in terms of EQ-5D-5L-based QALYs because it was associated with 0.036 more QALYs gained per participant, but this difference was not statistically significant (95% CI, -0.007 to 0.080;  $P = 0.098$ ). For the older cohort, the resulting ICER was estimated at \$15 709 per QALY gained (95% CI, -\$19 780 to \$51 199), which is below the \$50 000 per QALYs gained threshold for determining cost-effectiveness in Australia. Therefore, in the older cohort, the intervention was cost-effective (table 9).

Figure 1 presents a cost-effectiveness plane for the older cohort to visually represent the incremental differences in costs and health outcomes between the intervention and control groups. The cost-effectiveness plane suggests some uncertainty, with only approximately 35% of the bootstrapped paired estimates of mean differences in costs and QALY scores being in the south-east quadrant, which represents scenarios where the intervention is both less expensive and more effective than usual care.

Figure 2 shows this further in the cost-effectiveness acceptability curve, which shows that, at 12 months, the probability of the intervention being the cost-effective option compared with usual care at a WTP threshold of \$50 000 per QALY gained was approximately 74%.

**Table 9. Mean costs (per participant) and EQ-5D-5L outcomes**

Variable	Control		Intervention		Intervention effect <sup>A</sup>	
	Baseline	12 months	Baseline	12 months	Coefficient (95% CI)	P
Adults and older cohort						
Costs (\$)						
Hospital	5127 (414)	5925 (952)	3255 (394)	4781 (1030)	728 (−3641 to 5095)	0.74
Medicare	3116 (146)	3103 (125)	3238 (120)	3431 (122)	206 (−430 to 843)	0.52
PBS	2213 (134)	2496 (137)	2217 (292)	2778 (220)	277 (−453 to 1008)	0.46
Intervention	0	0	0	1000	1000 (1000 to 1000)	<0.001
Total costs	10525 (444)	11592 (896)	8739 (554)	12007 (1219)	2201 (−1765 to 6166)	0.28
Outcome						
EQ-5D-5L utility score	0.607 (0.015)	0.584 (0.015)	0.635 (0.014)	0.620 (0.014)	0.008 (−0.033 to 0.049)	0.71
QALYs gained <sup>B</sup>		0.595 (0.014)		0.627 (0.01)	0.032 (0.001 to 0.063)	0.042
ICER (\$)					69 585 (22 968 to 116 201)	
Adults cohort						
Costs (\$)						
Hospital	6284 (1253)	5917 (2543)	3250 (3181)	6604 (3044)	3721 (−7072 to 14513)	0.50
Medicare	2729 (177)	2292 (164)	2899 (221)	2937 (108)	475 (−488 to 1438)	0.33
PBS	1853 (43)	1865 (175)	1927 (443)	2311 (269)	372 (−647 to 1390)	0.48
Intervention	0	0	0	1000	1000 (1000 to 1000)	<0.001
Total costs	10 862 (1319)	10 068 (2559)	8053 (3489)	12 818 (3225)	5559 (6166 to 17 283)	0.35
Outcome						
EQ-5D-5L utility score	0.531 (0.029)	0.555 (0.026)	0.529 (0.027)	0.572 (0.026)	0.019 (−0.038 to 0.077)	0.51
QALYs gained <sup>B</sup>		0.543 (0.015)		0.550 (0.024)	0.007 (−0.068 to 0.082)	0.85
ICER (\$)					841 202 (170 447 to 1 852 850)	
Older cohort						
Costs (\$)						
Hospital	4424 (325)	5843 (655)	3355 (226)	4127 (685)	−647 (−2822 to 1528)	0.56
Medicare	3342 (121)	3562 (134)	3334 (161)	3589 (227)	−35 (−700 to 770)	0.93
PBS	2413 (226)	2847 (251)	2340 (194)	2970 (358)	197 (−304 to 697)	0.44
Intervention	0	0	0	1000	1000 (1000 to 1000)	<0.001
Total costs	10 175 (385)	12 254 (862)	9078 (278)	11 728 (1046)	571 (−2287 to 3428)	0.70
Outcome						
EQ-5D-5L utility score	0.644 (0.017)	0.598 (0.018)	0.677 (0.015)	0.639 (0.017)	0.008 (−0.031 to 0.047)	0.69
QALYs gained <sup>B</sup>		0.621 (0.018)		0.658 (0.008)	0.036 (−0.007 to 0.080)	0.10
ICER (\$)					15 709 (−19 780 to −51 199)	

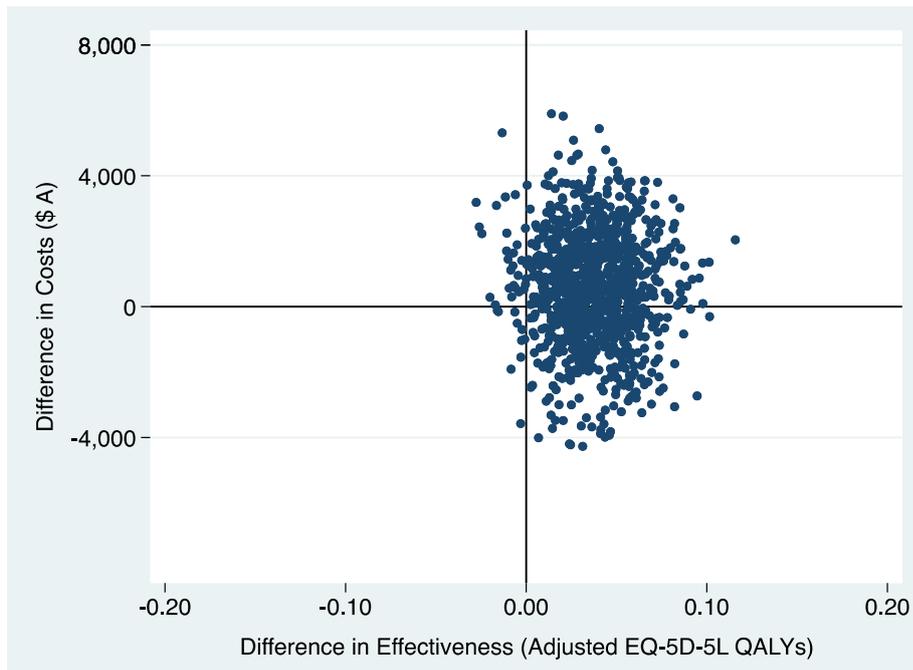
Unless indicated otherwise, data are presented as the mean (SEM).

<sup>A</sup> The intervention effect is calculated from a multilevel linear regression model.

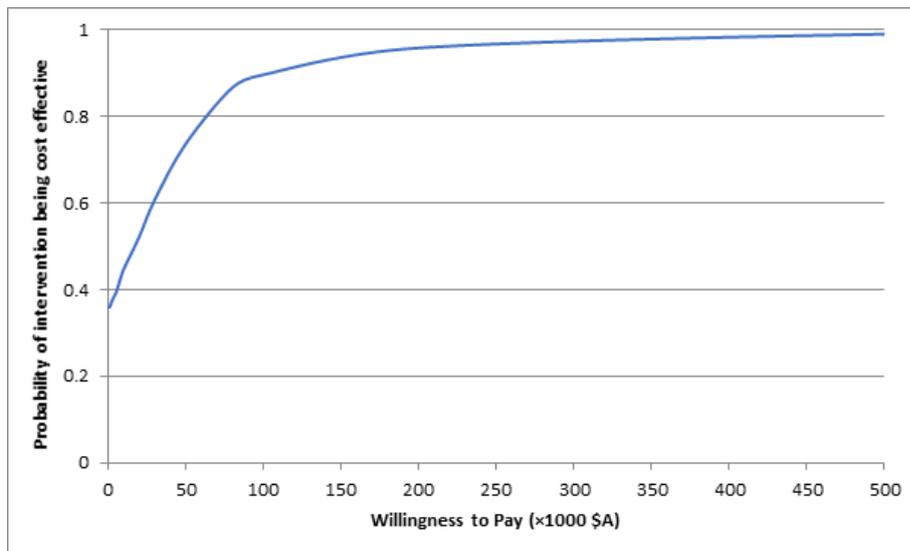
<sup>B</sup> QALYs were calculated from the EQ-5D-5L utility score<sup>15</sup>, and QALY gains were adjusted for differences in EQ-5D-5L scores between the two groups at baseline.

CI = confidence interval; ICER = incremental cost-effectiveness ratio; PBS = Pharmaceutical Benefits Scheme; QALY = quality-adjusted life-years.

**Figure 1. Cost-effectiveness plane: older cohort. QALYs, quality-adjusted life-years.**



**Figure 2 Cost-effectiveness acceptability curve: older cohort.**



*Threshold and sensitivity analyses*

If the incremental effectiveness of the intervention in the entire cohort remained unchanged 12 months after baseline (i.e. if it was still associated with QALY gains of 0.032 QALYs gains per participant), the intervention costs would have to be reduced by just over 60% (from \$1000 to approximately \$395 per participant) for the intervention to become cost-effective at the WTP threshold of \$50 000 per QALY gained. In the adults cohort, the intervention would still not be cost-effective even if the intervention cost were reduced to zero (i.e. the intervention would be associated with an ICER of \$619 703 per QALY gained, which is higher than the \$50 000 per QALY cost-effectiveness threshold). In the older cohort, reducing the intervention cost by any amount would only make the intervention more cost-effective. Further, the intervention would remain cost-effective at

this threshold even if the cost of the intervention were increased, provided this increase was not greater than \$1248 (i.e. from \$1000 to, at most, \$2248 per participant).

With respect to longer term costs and QALY gains associated with the intervention and usual care groups for the entire adults and older cohort, the intervention was associated with ICERs of less than \$50 000 per QALY when an assumption was made that the effectiveness of the intervention was maintained at 100% (\$11,961 [-\$73,931 to \$34,394]), 75% (\$14,283 [-\$88,283 to \$41,070]) and 50% (\$17,724 [-\$109,550 to \$50,964]) for up to 36 months after baseline. When the same assumption was made for the adults cohort, the intervention was associated with ICERs greater than the \$50 000 per QALY threshold, regardless of the time horizon (data not shown). This was the case even when costs and QALY gains were extrapolated to up to 10 years after baseline (data not shown). Making the same assumption for the older cohort at all time horizons led to the usual care group being dominated by the intervention (i.e. the intervention was both cheaper and more effective than usual care) (data not shown).

A sensitivity analysis was performed to assess the effect of including additional study costs on cost-effectiveness. These additional study costs covered the \$10 000 payment to practices for administrative costs associated with patient recruitment, data collection activities, taking part in qualitative interviews at the conclusion of the trial and practice facilitation by a research nurse estimated at 0.5 FTE.

For the adults and older cohorts combined, the intervention was still more expensive than usual care by \$2284 per participant, although this difference was not statistically significant (95% CI, -\$547 to \$5115;  $P = 0.11$ ). The resulting ICER was estimated at \$73 548 per QALY gained (95% CI, -\$35 114 to \$182 211), which is above the \$50 000 per QALYs gained threshold for determining cost-effectiveness in Australia.

For the adults cohort, the intervention was even more expensive than usual care by \$5642 per participant, although this difference was again not statistically significant (95% CI, -\$888 to \$12 171;  $P = 0.09$ ). The ICER for this cohort was estimated at \$949 429 per QALY gained (95% CI, \$896 189 to \$1 002 669), which is nearly 20 times the Australian QALY threshold for determining cost-effectiveness.

For the older cohort, the intervention was still more expensive than usual care, but by only \$654 per participant, and this difference was not statistically significant (95% CI, -\$2148 to \$3455;  $P = 0.65$ ). The ICER for this cohort was estimated at \$18 711 per QALY gained (95% CI, -\$21 319 to \$48 289), which is lower than the QALY threshold for determining cost-effectiveness in Australia.

## 7. Process evaluation results by age group

In the adults cohort, the continuity of care intervention effect showed statistically significant higher odds of GP appointments with the most frequently seen GP (odds ratio [OR], 1.32; 95% CI, 1.06–1.65;  $P = 0.013$ ). However, no statistically significant intervention effects were found for appointment length and follow-up after emergency department or hospital care episodes (table 10).

In the older adult cohort, the appointment length intervention effect showed statistically significant higher odds of longer appointments (OR, 1.37; 95% CI 1.17–1.59;  $P < 0.001$ ). However no statistically significant intervention effects were found for continuity of care and follow-up after emergency department or hospital care episodes (table 10).

**Table 10. Process evaluation: adults and older adult groups**

Indicator	Control		Intervention		Intervention effect <sup>A</sup>	
	Baseline	12 months	Baseline	12 months	OR (95% CI)	<i>P</i>
Adults cohort						
No. of participants	169		143			
Continuity of care (UPC) <sup>B</sup>					1.32 (1.06–1.65)	0.013
No	624 (29.1%)	618 (31.9%)	537 (29.4%)	471 (28.4%)		
Yes	1518 (70.9%)	1321 (68.1%)	1291 (70.6%)	1188 (71.6%)		
Appointment length					1.01 (0.81–1.27)	0.92
Brief + standard	1547 (72.2%)	1333 (68.8%)	1238 (67.7%)	1044 (62.9%)		
Long + prolonged	595 (27.8%)	606 (31.3%)	590 (32.3%)	615 (37.1%)		
Follow-up <sup>C</sup>					0.72 (0.29–1.81)	0.49
No	63 (61%)	41 (39%)	62 (52%)	58 (48%)		
Yes	51 (54%)	44 (46%)	32 (51%)	31 (49%)		
Older cohort						
No of participants	305		342			
Continuity of care (UPC) <sup>B</sup>					1.07 (0.93–1.24)	0.34
No	1162 (28.5%)	1216 (31.6%)	1203 (27.6%)	1233 (29.3%)		
Yes	2918 (71.5%)	2636 (68.4%)	3161 (72.4%)	2983 (70.8%)		
Appointment length					1.37 (1.17–1.59)	<0.001
Brief + standard	3022 (74.1%)	2839 (73.7%)	3066 (70.3%)	2726 (64.7%)		
Long + prolonged	1058 (25.9%)	1013 (26.3%)	1298 (29.7%)	1490 (35.3%)		
Follow-up <sup>C</sup>					1.38 (0.74–2.58)	0.32
No	75 (41.2%)	107 (58.8%)	111 (51.4%)	105 (48.6%)		
Yes	66 (52.0%)	61 (48.0%)	86 (54.4%)	72 (45.6%)		

Unless indicated otherwise, data are presented as *n* (%) of appointments.

<sup>A</sup> The intervention effect (odds ratio) was calculated from a multilevel logistic linear regression model.

<sup>B</sup> Data are *n* (%) of GP appointments with most frequently seen GP during the time-period.

<sup>C</sup> Data are *n* (%) of GP appointments within seven days of an ED presentation or hospital discharge.

‘Baseline’ refers to the 12-month period prior to intervention, ‘12 months’ refers to 12-month intervention period.

The analysis dataset was the adults ( $n = 312$ ) and older ( $n = 647$ ; excluding three participants with zero GP appointments at either baseline or the 12-month follow-up) Medicare matched datasets.

CI = confidence interval; OR = odds ratio; UPC = Usual Provider of Care index.

## 8. Case file review

A case file review was conducted in one large intervention group practice. In this practice, only modest improvements in continuity of care, appointment length and follow-up after a health event were observed despite a very systematic implementation of the intervention.

The case file review was performed after the intervention period had concluded using Medicare data and SA Health hospital records to enable cross-checking with practice records. The case file review focused on instances where participants had not received appointments with their preferred GPs, had not received longer appointments or had not been followed-up after an ED presentation or hospital discharge.

The primary reason participants did not have appointments with their preferred GP was that the appointment was for an acute or urgent problem (e.g. urinary tract infection or respiratory tract infection) and the preferred GP was not working on that day. Further appointments relating to this episode of care usually continued with the non-preferred GP until the problem was resolved.

Appointments with non-preferred GPs also occurred for vaccinations or wound care. These procedures were typically performed by a Practice Nurse with GP review. The GP review was typically performed by a duty doctor and not the participant's preferred GP. Several participants had consultations with a GP at the practice who specialised in skin checks. A small number of participants had GP procedures (e.g. venesection) performed at specialised centres.

Continuity of care at the practice level could be assessed from the case file review. In the 12 months prior to the intervention, 92.3% of GP appointments were with GPs at the practice. This increased during the 12 months of the intervention period to 95.4%. The reason(s) for the small (4.6%) number of appointments made with GPs who were not at the practice could not be established from the case file review.

Where appointments were with non-preferred GPs (at either the case review practice or another practice), invariably these were standard-length appointments. A small number of participants expressed a preference for more frequent standard-length appointments rather than longer appointments. However, there were occasions of standard-length appointments with a participant's preferred GP. In these cases, often the participant had used the emergency 'spare' QUEST appointment slot that had been blocked out (to promote continuity of care) as a standard-length appointment.

With regard to GP follow-up after a health event such as an ED presentation or hospital discharge, in a relatively small number of cases the practice did not receive the discharge summary and the participant did not inform the practice of their health event. Where discharge summaries had been received, the clinical team decided on occasions that a face-to-face appointment was not clinically warranted, and no further action was taken. On other occasions, the participant was followed up but by telephone and this telehealth consultation was not captured in our Medicare-defined measure of follow-up (i.e. face-to-face consultation). In a small number of cases, practice administrative staff failed to identify that the discharge summary related to a QUEST participant and the discharge summary was not brought to the attention of the clinical team.

## References

1. The Royal Australian College of General Practitioners. Standards for general practices. 5th ed. East Melbourne, Vic: RACGP, 2020. <https://www.racgp.org.au/running-a-practice/practice-standards/standards-5th-edition/standards-for-general-practices-5th-ed>. (viewed March 2022).
2. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; 41: 582–592.
3. McCaffrey N, Kaambwa B, Currow DC, et al. Health-related quality of life measured using the EQ-5D-5L: South Australian population norms. *Health Qual Life Outcomes* 2016; 14: 133.
4. Jayasinghe UW, Proudfoot J, Barton CA, et al. Quality of life of Australian chronically-ill adults: patient and practice characteristics matter. *Health Qual Life Outcomes* 2009; 7:50.
5. Reed RL, Roeger L, Howard S, et al. A self-management support program for older Australians with multiple chronic conditions: a randomised controlled trial. *Med J Aust* 2018; 208: 69–74.
6. Macran S, Brooks R, Rabin RE, de Charro FTH. Test re-test performance of EQ-5D. The measurement and valuation of health status using EQ-5D: a European perspective. Amsterdam: Kluwer Academic Press, 2003.
7. Hemming K, Marsh J. A menu-driven facility for sample-size calculations in cluster randomized controlled trials. *Stata Journal* 2013; 13: 114–135.
8. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) – explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health* 2013; 16: 231–250.
9. Glick HA, Doshi JA, Sonnad SS, et al. Economic evaluation in clinical trials. Oxford University Press, 2014. [eBook]2014 4/5/2021]. <https://ebookcentral.proquest.com/lib/kcl/detail.action?docID=1811851>.
10. Drummond M, Drummond M. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford, New York: Oxford University Press, 2005.
11. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; 20: 1727–1736.
12. Golicki D, Niewada M, Buczek J, et al. Validity of EQ-5D-5L in stroke. *Qual Life Res* 2015; 24: 845–850.
13. Nolan CM, Longworth L, Lord J, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. *Thorax* 2016; 71: 493–500.
14. Lu WS, Huang SL, Yang JF, et al. Convergent validity and responsiveness of the EQ-5D utility weights for stroke survivors. *J Rehabil Med* 2016; 48: 346–351.
15. Norman R, Cronin P, Viney R. A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. *Appl Health Econ Health Polic* 2013; 11: 287–298.
16. Kind P, Hardman G, Leese B. Measuring health status: information for primary care decision making. *Health Policy* 2005; 71: 303–313.

17. Kreimeier S, Oppe M, Ramos-Goni JM, et al. Valuation of EuroQol Five-Dimensional Questionnaire, Youth Version (EQ-5D-Y) and EuroQol Five-Dimensional Questionnaire, Three-Level Version (EQ-5D-3L) Health States: the impact of wording and perspective. *Value Health* 2018; 21: 1291-1298.
18. West BT, Welch KB, Galecki AT. *Linear mixed models: a practical guide using statistical software*. 2nd ed. Boca Raton: Chapman & Hall/CRC, 2014.
19. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005; 14: 487–496.
20. Barber JA, Thompson SG. Analysis and interpretation of cost data in randomised controlled trials: review of published studies. *BMJ* 1998; 317: 1195–1200.
21. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990; 10: 212–214.
22. Harris AH, Hill SR, Chin G, et al. The role of value for money in public insurance coverage decisions for drugs in Australia: a retrospective analysis 1994–2004. *Med Decis Making* 2008; 28: 713–722.
23. Pharmaceutical Benefits Advisory Committee. *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3)*. Canberra: Department of Health, 2016. <https://pbac.pbs.gov.au> (viewed March 2022).