Positive Changes Associated With a Recovery-Oriented Mental Health Care Training Intervention in the REFOCUS-PULSAR Specialist Care Cluster Stepped-Wedge Randomised Controlled Trial

Graham Meadows*, Lisa Brophy, Frances Shawyer, Joanne C. Enticott, Ellie Fossey, Christine D. Thornton, Penelope Weller, Elisabeth Wilson-Evered, Vrinda Edan, Mike Slade

Author details

- Graham Meadows MD, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia; Melbourne School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia; Monash Health, Melbourne, VIC, Australia.
- Lisa Brophy PhD, Centre for Mental Health, Melbourne School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia. Mind Australia, Heidelberg, VIC, Australia. School of Allied Health, La Trobe University.
- Frances Shawyer PhD, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia.
- Joanne C. Enticott PhD, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health; Department of General Practice, School of Primary and Allied Health Care, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia; Centre of Research Excellence in Suicide Prevention, Brain and Mind Centre, University of Sydney, Camperdown, NSW, Australia.
- Ellie Fossey PhD, Department of Occupational Therapy, School of Primary and Allied Health Care, Monash University - Peninsula Campus, Frankston, VIC, Australia.
- Christine D. Thornton BA, Ermha Ltd, Dandenong, VIC; Core Advanced Technologies Limited, Worcestershire, UK; Melbourne School of Law, University of Melbourne, Parkville, VIC, Australia.
- Penelope June Weller PhD, Graduate School of Business and Law, RMIT University, Melbourne, VIC, 3001 Australia.
- Elisabeth Wilson-Evered PhD, Institute of Health and Sport, Victoria University, Melbourne, VIC, Australia.
- Vrinda Edan, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia.
- Mike Slade PhD, School of Health Sciences, Institute of Mental Health, University of Nottingham, Triumph Road, Nottingham, NG7 2TU, UK.

*Corresponding author:
Professor Graham Meadows, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Dandenong Hospital, 126 - 128 Cleeland St, Dandenong Victoria 3175, Australia. Tel: (61 3) 9902 9696; Fax: (61 3) 9902 9900;
graham.meadows@monash.edu

This project was undertaken through Monash University.

Summary

Background: Recovery-oriented practice promotes individual strengths and recovery potential. PULSAR, adapting the UK-developed REFOCUS recovery-oriented staff intervention for Australian use, aimed to establish whether consumers accessing mental health services where staff had received the REFOCUS-PULSAR intervention showed increased recovery compared to consumers of non-intervention services.

Methods: A pragmatic two-step stepped-wedge randomised controlled trial at 18 sites grouped into 14 clusters across Public Mental Health Services (PMHS) and Mental Health Community Support Services (MHCSS). Staff training was refined between step-one and step-two. The primary (stream-one) outcome measure was the Questionnaire about the Process of Recovery (QPR) with cross-sectional data collected across three time-points. Stream-two, with two data-collection points, included five outcome-measures and five experience-measures. This trial is registered with ANZCTR, number ACTRN12614000957695.

Findings: Half of the available staff were trained (190), with substantial staff turnover across the three organisations (27-47%). Between 2014 and 2017, 942 stream-one consumer participants were recruited over three time-points (T0: 301; T1: 334; T2: 307) with 273 stream-two participants recruited at intervention-related time-points. (baseline: 140, follow-up: 133). The main mixed-effects model showed a small significant overall positive intervention stream-one effect of 3·7 (95% Confidence interval: 0·5– 6·8). Examining interactions, the mean difference between intervention and control groups at year-one also was 3·7 (95% Confidence interval: 0·6 – 6·8); findings were strongest for PMHS step-two. Stream-two findings of small effects, typically below study power threshold, favoured the intervention condition for all but one measure.

Interpretation: The REFOCUS-PULSAR intervention showed modest but distinct effectiveness in promoting recovery-oriented practice across sectors.

Funding: Victorian Government Mental Illness Research Fund.

Key words: Recovery, Recovery-oriented practice, Specialist mental health services, Mental Health, Training, Psychiatry, Cluster Randomised Controlled Trial, Pragmatic trial, Health services research, Complex intervention, Questionnaire about the Process of Recovery (QPR).

Research in context

Evidence before the study

Searching PsycINFO, Medline and CINAHL, for articles published in English between 1 January 2007 and 31 July 2017, given the development and evaluation of approaches to implementing recovery-oriented practices is relatively recent. The search strategy included the following search terms: [Mental Health/ OR "mental health" OR Mental Health Services/ OR "mental health service*"] AND [Recovery/ OR recover*], then identification of further relevant articles from reference lists of key papers, author searches and citation searches in Google Scholar. We selected articles if they were set in community mental health services and included data related to staff views, staff-related outcomes or consumer-related outcomes in the context of staff training in recovery-oriented practice (ROP) and/or implementation of ROP to promote and support personal recovery. This identified 16 relevant studies typically assessing staff-related outcomes after recovery-oriented training programs. While only REFOCUS had been evaluated using a randomised controlled trial design, these studies generally suggest that recovery-oriented training improved staff knowledge and attitudes towards recovery and improved self-efficacy towards providing recovery-oriented care, with a recurrent theme that the organisational culture of the service setting, and the provision of follow-up coaching appear to be important determinants of implementation success. Apart from the REFOCUS trial published in 2015, no others have reported whether consumer outcomes were improved by these interventions.

Added value of the study

The REFOCUS-PULSAR staff training intervention, adapted for Australian service settings from the REFOCUS package and based on the CHIME (Connectedness, Hope, Identity, Meaning, and Empowerment) conceptual framework of personal recovery, was examined through a stepped-wedge randomised controlled trial with quantitative assessment of effect on consumer-rated experience of recovery. Positive findings for intervention effect in the study provide evidence that the REFOCUS-PULSAR intervention as developed and implemented in this study brought about modest improvements in consumer-rated recovery for people using the involved services. The findings also suggest possible improvements in clinical recovery and experience of service.

Implications of all the available evidence

Training health-care workers to deliver recovery-oriented care using the REFOCUS materials developed over time and adapted to local settings can positively influence the process of personal recovery for consumers.

Introduction

Developing evidence around recovery orientation

The construct of recovery now commonly used in mental health care has roots in consumer perspectives[1] and may be distinguished from other conceptualisations by reference to *personal* rather than *clinical* recovery.[2] Recovery-oriented practice (ROP) involves clinical and other staff facilitating a change process through which individuals who have been diagnosed with mental illness are supported to live a self-directed life and to strive to reach their full potential.[3] Promoting recovery within mental health services is well established in mental health policy internationally[4] and in Australia[5] where this study is set. However, the practice lags behind policy: service-level intervention is required to effectively implement practices through which mental health professionals employ skills, values, attitudes and behaviours that support individuals in their personal recovery.[6] The past decade has seen the development of a number of recovery-oriented training programs, such as REFOCUS[6] and THRIVE[7] in the UK, the Collaborative Recovery Model[8,9] in Australia and Person-Centred Recovery Planning[10] in the USA. They typically emphasize the use of coaching and person-centred, strengths focused and collaborative processes for supporting service users in their recovery. A useful reference framework for the work on training interventions may be Kirkpatrick's four levels of learning evaluation: K1-reaction, K2-learning, K3-behaviour and K4-results.[11] The literature is strongest on levels 1 and 2, with few programs having evidence at either level 3 or level 4. Typically work at level 4 has not had the strength of evidential value that goes with RCT methods so there is a need for further evidence at this level. Evidence of the effectiveness of these interventions to promote ROP is required across settings, so that they might be adopted with some confidence by services working towards these policy goals.

From REFOCUS to PULSAR – a developmental trajectory

REFOCUS is a staff training intervention developed and trialled in the UK.[2,6,12] In a developmental process informed by the theory of planned behaviour,[13] working towards changing both what practitioners might do with consumers of mental health services (consumers) and how they might do it,[14] the REFOCUS intervention came to include, as elements of a team-based training intervention

for community mental health teams in England, three working practices of 'understanding values and treatment practices', 'working to strengths', and 'supporting goal striving'. So, the REFOCUS intervention was designed to promote recovery through changes in staff and team skills, knowledge, behaviour, values, and relationships with consumers.[2,12]

In a large-scale cluster Randomised Controlled Trial (cRCT), outcomes of usual care plus REFOCUS were compared with usual care only in 27 community mental health teams delivering services to adult consumers with psychotic disorders. In primary analyses, personal recovery assessed using the consumer-rated Questionnaire about the Process of Recovery (QPR)[15] did not differ between the intervention group and controls. Secondary analyses suggested higher team-participation was associated with higher staff-reported recovery-promoting behaviour and improved QPR. Possible reasons advanced for the negative primary analyses that might be modifiable in subsequent work included the following issues:[6]

1. The REFOCUS recruitment protocol and criteria meant that, on average, consumer participants had been using mental health services for >15 years, suggesting the possibility of entrenched ways of relating to services, and problems that may take longer than one year to change.
2. Participant attrition, higher than anticipated in this 12-month longitudinal study (26% vs 7%), resulting in a reduction in planned statistical power.
3. Inclusion of adaptive design principles[16,17] might be advantageous.
4. Future designs might either use a homogenous team-type or stratification by team characteristics.
5. Transition to ROP might require organisation-wide rather than team-level strategies.

The 'Principles Unite Local Services Assisting Recovery' (PULSAR) work program was based in Victoria, Australia. The REFOCUS team advised on project development enabling PULSAR, four years behind REFOCUS in development and implementation, to benefit from lessons learned during REFOCUS. Changes to the intervention included adjustments to the REFOCUS materials to enhance relevance to the local setting and to incorporate developments made in the course of the REFOCUS work after the REFOCUS manual[18] was concluded for study use. The intervention here is referred to as "REFOCUS-PULSAR" (shortened to "PULSAR" in the protocol paper and local implementation[19]) since while it was developed for the PULSAR study,[19] it drew heavily on REFOCUS materials.

The research approach,[19] chosen based on addressing issues 1-5 above, involved adoption of a specific cRCT variant involving Stepped-Wedge intervention allocation (a cRCT-SW) where all study sites receive the intervention but time of intervention is allocated randomly, here according to

two ‘steps’, step-one and step-two. Since those people who may benefit most from ROP in relation to personal recovery may also experience clinical recovery and so be discharged earlier from treating services, sampling based on people with long-term service tenure may bias against positive findings as noted in point 1 above. Hence, the PULSAR design primary recruitment strategy recruited independently at three time-points (baseline: T0; year 1: T1; year 2: T2) with tight control on consistency of recruitment processes so that sampling bias is minimised as a source of systematic error in findings related to intervention effect across time-points. The cRCT-SW research design with repeated cross-sectional recruitment, then, carried possible advantages for point 1-2 above. The two-year two-step stratified cRCT-SW approach promised greater possibility for progressive refinement of the training intervention through experience, providing some response to point 3. Randomisation in this study was stratified by team type addressing point 4 above. The design also went a small way to address point 5 above since in the later stage of the stepped-wedge design the implementation was in effect organisation-wide across community services.

Aims and hypotheses

The aim of this pragmatic cluster stepped-wedge randomised controlled trial was to evaluate the effectiveness of the REFOCUS-PULSAR staff ROP training intervention for improving the experience of personal recovery as reported by consumers using repeated cross-sectional samples. The primary hypothesis was that consumers in the REFOCUS-PULSAR post-intervention clusters would experience significantly greater personal recovery compared to consumers accessing other mental health services that at relevant time-points within the cRCT-SW had not received the intervention. We also investigated change in clinical recovery and experience of the services.

Methods

Setting

Participating services were providers of mental health care to people living in the catchment area of a large Public Mental Health Service (PMHS) in Victoria, Australia. The area ranges from a relatively affluent coastal city area to the most socio-economically disadvantaged and culturally-diverse area in metropolitan Melbourne and includes a semi-rural growth-corridor. In Victoria, state-run area-based and block-funded PMHSs, typically accessed by people with more severe mental illnesses, include clinical services comprising a range of teams and service types. Here are included inpatient units, community-based residential rehabilitation, continuing care, and community treatment teams. Acute or longer-term Residential care is typically provided in units of around 25 beds. Caseloads in

community services vary from around 10 in Mobile Support and Treatment Services (MSTS) to 25-35 in many community clinics while typical length of care with a particular team may vary between a few days with Crisis Assessment and Treatment Teams (CATTS) to several years with MSTS and Community Care Units (CCUs). Mental health care funded by the Victorian government also includes substantial investment in the Mental Health Community Support Services (MHCSS) sector which, run by non-government organisations, provides residential and outreach psychosocial support. Within this setting, the temporal context for the work through 2014-2016 included events worthy of some comment – details on these are provided in Appendix 1).

The State-funded organisations that operated in the catchment were the major PHMS and two organisations from the MHCSS sector. Specialist care sites or teams within these organisations were identified by the PMHS and MHCSS service partners then approached; all agreed to participate.

Design overview

Specialist-care PULSAR project data collection from consumers included three streams. Stream-one, a cross-sectional complete step-wedge cRCT with self-administered instruments, collected QPR and demographic data. The QPR, identified as the primary outcome[19] was the basis for stream-one power calculations. Stream-two, a cross-sectional pre- and post-intervention incomplete step-wedge cRCT, involved face-to-face interviews with a subset of stream-one participants. Stream-three, a longitudinal incomplete step-wedge cRCT involving consumers from Stream-two with diagnosed psychotic disorders, did not achieve adequate recruitment targets and is not reported here.

Participants

Staff

Participating teams' members were eligible to receive the PULSAR training intervention if they were working part-time or full-time in a direct service role and had an active caseload with consumers being recruited for the evaluation. Casually employed staff or those also working in a non-intervention site at the time of training were ineligible.[19]

Consumers Stream-one

Eligible consumers were: receiving care from a participating cluster with contact in the three months prior to data collection; aged 18-75; able to provide informed consent; proficient in English; and not imprisoned. Eligibility screening, conducted by administration and clinical staff at participating organisations, used detailed instructions provided by the research team. A letter sent to all eligible consumers from participating sites invited completion and return of a demographics/QPR survey form and a contact details/consent to be contacted for a face-to-face interview form. An AUD$10

shopping voucher was sent to participants for returned surveys where contact details were provided. Additional recruitment strategies to encourage consumer response to the mailouts were utilized according to site need. Strategies included, for example, having researchers, including consumer researchers, speak about PULSAR at participating sites and use of PULSAR-branded publicity materials.[19] Through an active quality assurance process monitoring recruitment, and because this was important to the design, the balance of recruitment between onsite recruitment and mailout approaches was kept as consistent as possible across timepoints and clusters. Decisions on whether or not to repeat bulk mailouts for given clusters or continue onsite recruitment were based on a weekly review of QPR numbers by recruitment method by cluster and taking into consideration the need to also recruit sufficient numbers for face-to-face interviews. Time spent recruiting at T1 and T2 at a given cluster was matched to T0 activity at the same cluster and only adjusted if necessary to match the number of QPRs collected via this method.

Consumers stream-two

Consumers were eligible for stream-two and recruited by phone, email or letter if they had provided contact details, consent to be contacted for this purpose and were at the pre- or post-phase of an active intervention site at the time of recruitment.

Randomisation and masking

Eighteen care-delivery teams, grouped into 14 clusters to enable adequate recruitment in the context of some smaller teams, were classified into seven strata. Team characteristics varied so strata groupings included teams similar in specified function. Within PMHS these were: CATTS (x3 teams; two smaller teams grouped into one cluster) and MSTS (x2 teams); CCUs (x2; grouped with MSTS, being smaller teams and introduced earlier as having shared focus on long term intensive work with people with more complex needs); Community Mental Health Services/Continuing Care Teams (x4). The remaining stratum included services delivered by two participating MHCSS, here designated MHCSS-1 and MHCSS-2. These were: Prevention and Recovery Care services (PARCs; x4) delivering short-term, subacute, residential recovery-oriented care; and Community Outreach Services (x3; two from the one organisation grouped into one cluster).

Stratified randomisation was used to allocate clusters to receive the intervention in either step-one or step-two using an online Research Randomiser with randomisation keys corresponding to the seven strata and allocation of clusters within strata to step-one or step-two in the cRCT-SW design. Randomisation was performed offsite by an independent researcher during the third quarter of 2014. As the intervention involves training, specialist mental health care staff knew their allocated condition as the study progressed. Consumer participants, however, were not informed if staff at their

service received the training and efforts were made to maintain the blindness of research assistants for onsite recruitment and stream-two interviews with consumers. Further details are in the protocol paper.[19]

Procedures

Intervention

The REFOCUS intervention[2,18] introduced earlier as developed in the UK to promote ROP is described in essence in a freely available manual.[18] The REFOCUS-PULSAR intervention comprises a manual[20] adapted from REFOCUS, a structured training intervention to support use of the REFOCUS-PULSAR manual, and follow-up sessions called PULSAR Active Learning Sessions (PALS).

REFOCUS-PULSAR development, following Medical Research Council Guidelines for Complex Interventions,[21] and the plan-do-study-act (PDSA) model as a method for controlling and improving process[17] was guided by discussions with the REFOCUS research team, consideration by a Lived Experience Advisory Panel (LEAP), and information from qualitative analysis of group sessions with staff from participating organisations. The content of the REFOCUS manual was substantially retained in the REFOCUS-PULSAR manual[20] with some amendments to contextualise it for the PULSAR study setting including legal and policy contexts. Additions - less than 25% of the manual - included material related to relapse-signatures and relapse-drills, and material on the CHIME ROP conceptual framework 'Connectedness, Hope, Identity, Meaning, and Empowerment'[12] which was developed during the course of the REFOCUS study. In summary, the REFOCUS-PULSAR intervention was grounded in experience and learning from REFOCUS, research evidence, government policy and law.

The REFOCUS-PULSAR training was supported by slide-presentations, a manual, session-plans and videos. In a change from the REFOCUS intervention, training was co-facilitated throughout by professional staff and trainers with lived experience of mental health problems, including the project's consumer researcher. This, based on local consultations, was expected to enhance the recovery-orientation of the training. Carer input featured in specific sessions. Quality assurance is described in Appendix 2.

The step-one intervention for clinical services was designed as a two-day session, with the community services training planned as a separate two-day session during the same week. In addition to having two project-employed consumer trainers, trainers were accessed from clinical services for clinical sessions and from the community sector for community sessions. This enabled the inclusion of specialist skills and experience in training delivery.[19] Step-two training was modified based on

analyses of participant and trainer evaluations from step-one. Details of changes can be found in Appendix 2. PALS, offered monthly as hour-long sessions to staff and managers of involved teams to support practice-based implementation of ROP, were facilitated by PULSAR investigators and local trainers.

Control condition

Standard treatment as delivered through the range of teams introduced above, was governed by national standards,[22] adherence to which is maintained by regular accreditation. Consumers of the service often will have their locus of care change in response to changing needs between the more intensive community teams (CATTS, MSTS), residential options including the PARCs, or less intensive community options. Case management in community clinics often functions to coordinate transitions through these levels of care and seeks to ensure that needs for medication, monitoring, supportive, and psychosocial interventions are met. Teams typically have multidisciplinary representation from mental health care disciplines with nursing as the largest single workforce component.

Adverse events

Anticipated possible study-related adverse events included: 1) risk of distress by a participant during an interview; 2) issues related to disclosure of potential self-harm or harm to others 3) risk of harm to staff. A risk-prevention and management protocol was approved by the governing HREC. Participants were provided with written contact details of the manager of the governing HREC for complaints. We did not systematically collect other adverse event information from consumers. For further details of adverse events and complaints procedures see Appendix 3.

Outcomes

Measures

These are divided into "Outcome Measures" (OMs), assessing clinical and personal recovery, and "Experience Measures" (EMs), assessing consumers' experience of health care. The primary stream-one OM (see Table 4, protocol paper[19]) was the QPR, a 22-item consumer-rated questionnaire used to assess personal recovery with each item being rated on a 5-point Likert scale ranging from 0 (disagree strongly) to 4 (agree strongly) and higher score indicating increased recovery.[15] While a 15-item scale has been suggested as a perhaps more robust alternative, this has not been independently validated other than within the 22 item questionnaire.[23] Having collected the 22-item version and powered the study based on known properties of this, we retain consistency with our protocol paper and focus on the 22 item score. In this study, Cronbach's alpha was 0·95 for both versions.

Secondary measures in stream-two, both consumer-rated were:

- EM: The importance of services in recovery questionnaire (INSPIRE) assessing recovery support from a worker[24] has sub-scales of support (20 items) and relationship with worker (7 items) scored by converting the mean of 5-point Likert ratings to a percentage.[24]
- OM: The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) assessing emotional and functional well-being has 14 Likert-scaled items with higher scores indicating greater mental well-being.[19]

Additional measures administered to consumers in stream-two (grouped as OMs and EMs) and reported here include:

EMs

- The Perceived Need for Care Questionnaire (PNCQ) assesses perceptions of mental health care, classifying consumer-identified perceived needs as unmet, partially met or met.[25]
- The Client Satisfaction Questionnaire (CSQ) assesses satisfaction with services.[26]
- The Mind Australia Satisfaction Survey (MASS) rates satisfaction with services, staff-consumer service delivery partnerships, and individual service-use outcomes.[27]
- The Coercion Ladder, a visual analogue scale, measures consumers' perception of coercion in mental health service interactions.[28]

OMs

- The Global Assessment of Functioning Scale (GAF) is a researcher-rated (0-100) positively rated measure of individual social, occupational and psychological functioning.[29]
- The Social and Occupational Functioning Assessment Scale (SOFAS), researcher-rated, (0-100) measures function independently from psychological condition severity.[29]
- Days out of role. This measures the impact of mental health problems on usual daily activities over the previous 30 days.

Participant demographic information was also collected.

Consent and key data collection timepoints

In stream-one, consent was by return of a completed survey. Stream-two participants provided written informed consent; interviews took around 60-90 minutes - interviewer blindness was assessed at completion (see protocol paper[19] for further details).

Baseline (T0) data collection occurred in the year prior to and three months after the delivery of the step-one intervention. The first three months after intervention delivery is deemed suitable for

baseline data collection based on the Kirkpatrick training evaluation model,[11] whereby the embedding of practice change is considered to take at least 9 months: 3 months for consolidation and 6 months for implementation. During both T1 and T2 periods, data collection at clusters sites took place at a minimum of 9 months after delivery of the intervention to allow embedding of intervention principles and practices.[19]

Staff finishing REFOCUS-PULSAR training were asked to complete a training evaluation (K1[11]) rating satisfaction from 1 "extremely dissatisfied" to 10 "extremely satisfied". Team managers or administrators were asked to record staff movements every three months.[19] The percentage of the team that attended at least one training session, in both headcount and full-time equivalent (FTE), was calculated for time of training. Team staff turnover was the percentage of staff who left, joined, or moved internally in the organisation but out of the cluster calculated on headcount.

Statistical analysis

Power

These calculations, using the sample size and power calculations described by Hemming and Girling[30] via Stata stepped-wedge V.11[31] were based on: 14 clusters; an intra-cluster correlation coefficient (ICC) of 0·05; significance level 0·05; power 0·80; and published standard deviations.[19] Stream-one and stream-two were powered for medium primary-outcome (QPR) effects. Stream-one detection of a change in mean QPR score by 6·34 indicated 756 surveys (252 in each wave, 18 per cluster per wave). Stream-two detection of a change in mean QPR score by 7.68, indicated 252 surveys (63 at baseline, 126 at step-one and 63 at step-two, 9 per relevant cluster per step). For stream-two secondary outcomes, expected detection thresholds were mean changes in WEMWBS of 4.8 and INSPIRE of 7.72 (medium effects).

Analysis plan

Intention-to-treat analysis was performed in line with a pre-specified analysis plan for all outcomes, using Stata (version 15). Participants were analysed in the groups to which their participating clusters were allocated. We analysed all outcomes using multi-level regression models (linear or Poisson regression as appropriate), with timepoint and intervention status as fixed effects, and clusters as a random effect. Timepoint was included as a categorical variable. Covariates, selected on statistical and clinical considerations, were age-group, gender, sector (PMHS/MHCSS) and step group (stream-one models only). No other covariates have yet been investigated for inclusion into the models, and a later separate investigation will explore the large pool of covariates and their effects on the study outcomes. Covariates of age-group and gender were included as they commonly influence clinical outcomes. Sector (PMHS/MHCSS) was included, as the most important stratification variable, but

not the other seven strata as this would have produced an overfitted model. Stream-one models included step group (step-one or step-two) - important temporal changes in the setting and changes in the intervention between steps are detailed in supporting materials. Step group could not be included in stream-two models due to collinearity with intervention status in the incomplete cRCT-SW design. It was anticipated (see protocol paper,[19]) that consumers would be modelled as random to account for repeated measures, but stream-one and stream-two cross-sectional recruitment attracted predominantly singletons, contributing to one timepoint only. Simulation studies have found low levels of bias for models with up to 70% singletons and 50 to 500 clustering units[32] so an adjustment to the analysis plan specified that participants would be specified as random only if less than 70% of data came from singletons.

Intervention effects are estimated from the models described above, recommended by Hussey and Hughes.[33] Also investigated and supplied as supplementary analyses in appendices are models with interaction effects between timepoint and intervention status, in which trends across the defined sector (PMHS and MHCSS) are reported.[34] The statistician was not blind to treatment allocation during the analyses.

Role of funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit it for publication.

Ethics approval

Approval was obtained from Monash Health (14102B) and Monash University (CF14/1600 – 2014000773) Human Research Ethics Committees.

Results

Implementation

Training and PALS

Step-one REFOCUS-PULSAR ROP training was delivered to 84 staff from the three services in the first quarter of 2015, in 22 days of workshops delivered by 7 trainers. Step-two training was held in June-July (plus an extra session in October) 2016 and delivered to 106 staff over 21 days by 8 trainers. In total 190 staff (111 PMHS; 79 MHCSS) were trained. On average across clusters, 49·1% (PMHS: 38·2%; MHCSS: 63·8%) of staff employed at the time of training attended at least one

training session. Adjusted for FTE, this was 51·2% (PMHS: 38·8%; MHCSS: 62·4%). Staff turnover was 42·2% for PMHS, 46·7% for MHCSS-1 and 26·7% for MHCSS-2.

Positively-judged training satisfaction (K1; scores >5) improved significantly from Year 1 to Year 2, Odds Ratio 2·71 (95% CI: 1·04, 7·05, p = 0·04). Staff trained included representatives of multiple disciplines but the team-based training approach in the most part did not succeed in engaging senior medical staff – it became apparent through the project that they more typically attend service-wide profession-specific trainings which would not readily be compatible with the cRCT model. A medical-specific training of 2 x 1·5 hour sessions was attended by 11 registrars but no consultants. For two PMHS teams no PALS occurred for logistical and engagement reasons. For all PMHS teams where they did occur (seven team settings including some that were combined), the mean total number of sessions was 8·1, SD 4·7. For 22% of these sessions, arranging team sessions was not successful so meetings were with individual clinicians. In MHCSS settings PALS came to be integrated into monthly staff support sessions and so the element of this that was PALS-specific cannot be quantified.

Consumer Recruitment

Between 18 September 2014 and 19 May 2017, 942 consumer participants were recruited across the three time-points, 575 from PMHS and 367 from MHCSS. Of these, 273 participants were recruited for stream-two interviews at timepoints related to the intervention delivery (baseline: 140, follow-up: 133). Overall recruitment targets were surpassed at each time-point (T0, T1 and T2) and most clusters were recruited into as planned (N=18 per cluster) at each time-point (see Figure 1, Figure 2, Appendix 4 and Table 1). As expected, overall recruitment rate from mailouts was low at 8·1% but yielding 622 or 66% of QPRs. Overall onsite recruitment rate as a proportion of all participants was 39·9% yielding 320 (34%) of all QPRs. Percentages of QPRs derived from onsite recruitment were 32% at T0, 34% at T1 and 36% at T2. Table 1 describes each cluster including: organisation sector, stratification level, allocated intervention step, and number of consumer participants recruited at each timepoint in both stream-one and stream-two. Table 2 shows the consumer descriptions in stream-one with further details in Appendix 5 which also included details of consumers recruited into stream-two.

Blindness

This was systematically assessed – see Appendix 6. We see it as unlikely through the course of the project that unblinding represented a significant bias to findings.

Outcomes

Model specification

In line with the adapted analysis plan, since in stream-one and stream-two 90% of the data arose from >50 (854, 254) singletons, consumer was not specified as random.

Primary outcome

The main model outputs in Table 3 show that, after adjusting for age, gender and step group and accounting for clustering, we find significant intervention and sector effects. The processes done to build the main model are in Appendix 7 and for the interaction term model in Appendix 8. Figure 3 presents the adjusted primary outcome means determined by the interaction term model.

Table 4 shows intervention effects, estimated as the difference in model-adjusted means (Table 3) between control and intervention data. This was 3·7 (95% Confidence interval: 0.5 – 6.8) for the primary outcome in stream-one, which was significantly greater than zero. To illustrate the degree of the effect size, and while there are some complexities in interpreting this in the context of the specific modelling, we have estimated Cohen's d for the intervention effect as the model adjusted difference (3·7) divided by the sample standard deviation (16·2) = 0·23 , which is a small effect.

Appendix 8 shows the model when including interaction terms, and show the overall mean difference between treatment and control groups at year 1 (model 1.6) was 3.7 (95% Confidence interval: 0.6 – 6.8) which was significantly greater than zero.

Figure 3 shows QPR scores over time by sector. Pre/post intervention differences occur between T0 and T1 for step-one clusters, and between T1 and T2 for step-two clusters. Therefore, four pre/post intervention scenarios are depicted in this figure (two in each sector). Two of these showed evidence of a significant pre/post intervention difference in QPR scores: in the PMHS sector (2a), in the step-two group there was a significant difference between T1 and T2 of 4·9 (z-score=3·0, p=0·003; Cohen's d estimate = 0.30, small-to-medium effect); and in the MHCSS sector (2b), in the step-one group there was a significant difference between T0 and T1 of 1·1 (z-score=2·7, p=0·006, Cohen's d = 0.07, small effect).

Secondary and other outcomes

Ten sets of results from stream-two are shown in Table 1. Analysis of findings from the PNCQ and a conclusion regarding direction of change are presented in Appendix 9. While none of the findings in Table 5 are individually statistically significant, for nine of ten analyses, central estimates suggested a mean change in the direction favouring the intervention, with estimated effect below the level of change for which the study was powered. If the intervention had no effect, then the probability of each result having direction favouring the intervention is 0·5 and the binomial probability that this

would occur nine times from ten results is 0·0107. So the findings suggest some modest positive influences across the span of these variables.

Discussion

Summary and interpretation of key findings

The project found a small but statistically significant effect on consumer stream-one QPR scores for the REFOCUS-PULSAR staff training intervention, involving two service sectors and delivered in context of a stepped-wedge design. Small effects in pragmatic trials are expected, and the significant finding is encouraging [35]. A significant interaction effect found for service sector suggests changes in sectors are better considered separately: In PMHS, there was no significant change from T0 to T1 for the step-one group - when this might have been expected because this was an intervention period. For the step-two group there was significant improvement from T1 to T2 (4.9 point increase in QPR scores) in their intervention period. In MHCSS, there was small but significant change (1·1) in step-one clusters through their intervention period (T0-T1) and a positive, though not significant, trend in step-two clusters from T1 to T2.

The 3·7 point improvement in QPR score represents a 5.7% change in the full scale score. Recommendations regarding the modelling approach used are that standardized effect sizes are easily distorted by factors unrelated to size of effect[36] and are not straightforward to interpret due to expected variance differences in the mixed model components.[37] Nevertheless the indicative calculation given of Cohen's d suggests a small positive effect so this is how we have framed our discussion. Based on QPR questionnaire content, changes of 1-2 points might be clinically meaningful. For instance a 2 point shift is achieved if the item 'I feel part of society rather than isolated' goes from neutral to strongly agree, which might represent a significant recovery outcome. The training team, working in a PDSA approach, made modifications to the training as delivered in step-two following feedback from step-one. These results seem to confirm that these modifications achieved an enhanced impact in PMHS step-two.

While speculative, mechanisms that might have led to greater primary outcome effect in step-two in PMHS might be that the attention given to the relationship between the two trainers (see Appendix 3) had the intended effect of providing better modelling of behaviour for participants through more clearly demonstrating respect for a lived experience perspective and more advanced communication skills. This perhaps also with introduction of dedicated content on coaching. Earlier availability of the manual may have improved uptake of principles for some participants while the team may also generally have gained experience with the delivery of both the core training and the PALS through

time. MHCSS findings may be influenced by pressures building in that sector through the course of the project as noted earlier and particularly potentially negatively influencing step-two findings. Stream-two findings included non-significant small effects, typically below the study power threshold. While conclusions here must be qualified, in nine of ten instrument comparisons the direction of central estimate of effect was in the direction favouring the intervention condition, a finding unlikely to be due to chance. At least it seems unlikely that any improvements in ROP came at a systematic cost in terms of other impacts. On balance of probabilities it is more likely that there was some small level of clinical and other benefit from the intervention.

Comparisons with REFOCUS

Findings here are more positive overall than those from the REFOCUS study. The differences developed between PULSAR and REFOCUS including those based on learnings from the REFOCUS experience may have influenced this. The literature on stepped-wedge designs had advanced in the period between design of REFOCUS and PULSAR and the adaptive nature of the PULSAR design allowed for refinements of the training following the first implementation to be evaluated. We note that if this study had been conducted with a similar parallel-group RCT design to that of REFOCUS, then without the inclusion of the step-two findings, PULSAR would not have yielded the positive findings reported here. The involvement of facilitators with lived experience of mental health issues and recovering is central to challenging conventional practices, and in making progress toward an effective recovery-oriented mental health workforce.[7] This might be why we achieved significant finding particularly in step-two PMHS, when the interaction between co-facilitators had been further developed.

Limitations

Accuracy of change-estimates might have been affected by the challenges facing the services as noted in the introduction. In both sectors the trend from T0 to T1 in the step-two group receiving no intervention in this time was of declining QPR scores, this most strongly in the MHCSS. Taking into account the challenging influences on all involved services, particularly MHCSS as noted in the 'Setting' section earlier and Appendix 1, it may be that these were acting across the services to drive QPR scores down. If that effect were also operating in the teams at the time they were receiving the intervention, then the underlying trend there might have been towards declining QPR as well. In this case, the findings might be underestimating the effect of the REFOCUS-PULSAR intervention.

REFOCUS-PULSAR training only managed to reach half of staff in intervention sites and few medical staff, which may have reduced intervention potency. In implementation outside constraints of a team-randomised cRCT, better results might be expected from greater engagement of medical staff whether in team-based or profession-specific training.

The REFOCUS intervention recommends some record-structure changes to support ROP, not possible in this cRCT because of organisation-wide regulation of form structures. In the PMHS since PULSAR concluded, the CHIME framework[14] has been integrated as a prompt into an organisation-wide record suite revision which has contributed to further interest in REFOCUS-PULSAR training.

Our recruitment strategy including repeated sampling and direct consumer approaches was chosen for strengths of avoiding clinician discretion as a key action-point for selection bias, enhancing consumer autonomy in participation,[38] and of avoiding selection bias towards greater chronicity of course of illness, identified as a problem in REFOCUS. However, while we have documented the considerable efforts made towards consistency of recruitment strategies, the possibility that this created time-variant selection bias on findings cannot be entirely excluded.

Further work

Policy on ROP has been described as "substantially ahead of research and practice"[6]; this is a values-based movement and policy and societal imperatives are strong that something be done to encourage services to work towards ROP even while evidence as to what is best to do may be accumulating. Multiple other ROP based trainings are in use with limited evidence at K1-3 and typically none at K4. The REFOCUS-PULSAR program can be considered for use based on reported findings suggesting improvements in high adopting teams in the English study, along with these K4 findings from PULSAR. There have been requests from teams in the participating PMHS for further PULSAR training, with exploration of extending and adapting the training to include inpatient staff so that the recovery-oriented culture can extend more widely across the care spectrum. In responding to these requests this team are mindful of the need to continue carefully to evaluate such initiatives, continuing PDSA cycles also with attention to educational evaluation at levels K1-K4 wherever possible and development of fidelity measures.

To better understand how sustained practice change can be achieved within services, future ROP training initiatives are recommended to strengthen the focus on implementation strategies, such as follow up coaching or mentoring, refresher programs, and service user feedback and evaluation.[8,10,39] Wide-ranging organisational factors are recognized as influential in supporting or constraining ROP implementation efforts,[9,10] so that attention to organisational readiness for change and alignment of

organisational policies, processes, staffing and resources with recovery oriented principles are also important. cRCT designs studying teams impede use of organisation-wide strategies and RCTs where randomisation is by organisation have limitations of large clusters so design considerations continue to be a challenge in accumulation of highest level evidence for these approaches.

Conclusions

Taken together, these results suggest that the REFOCUS-PULSAR intervention can lead to a modest overall measured improvement in personal recovery, also possibly with a small effect on some measures of clinical recovery and other aspects of client experience.[40] From an educational intervention perspective they place the REFOCUS-PULSAR intervention in the situation of having at least some evidence at level K4,[11] something otherwise lacking in the literature surveyed to date. It seems at very least unlikely that any improvement in ROP came at a cost in terms of clinical measures. While the findings of this study are modest, this is not surprising in a pragmatic trial and they provide at least some indication of positive change for consumers accessing the intervention services.

Contributors

GM was the Principal Investigator on this trial and together with JE led development of key elements of the design and analysis approach and interpretation of the findings. MS developed the original REFOCUS intervention and advised on adaptation. LB chaired the research module task-group and provided oversight to development and implementation of all elements of the design. FS provided overall coordination for field work and staff training and was centrally involved in the day-to-day operations of trial implementation. JE conducted the analyses. Specialist contributions regarding design elements were made by EF and EWE. CT contributed to the study design and implementation within MHCSS and was a chair of the adaptation module task-group. PW chaired the implementation module task-group which oversaw the delivery of the training intervention. VE, LB, GM, PW and EWE developed the specific specialist care training intervention and associated resources. GM, LB, FS, VE and EWE developed the specialist care instrumentation and fieldwork trial protocols. The core drafting group for this paper comprised GM, LB, FS, JE, MS; the remaining authors revised it critically for important intellectual content. All authors read and gave final approval for this version of the paper to be published.

Declaration of interests

We declare no competing interests.

Data sharing statement

In compliance with the requirements of the Monash Health Research Ethics Committee, the data supporting our findings in the manuscript will not be shared because we did not obtain participant consent to do so.

Acknowledgements

We thank: the partner organisations of ermha, Mind Australia and Monash Health for support to the project; the Steering Group and Lived Experience Advisory Panel for sharing their expertise; the REFOCUS-PULSAR trainers; and the team of PULSAR research and fieldwork staff for data collection. We thank all project participants who in sharing their time and experience made this work possible. This article presents independent research funded by the Victorian Government through its Mental Illness Research Fund. MS is supported by the NIHR Nottingham Biomedical Research Centre. The views expressed in this publication are those of the authors and do not necessarily reflect those of the Victorian Government including the Department of Health and Human Services. Further information is available at www.pulsarrecovery.org.au.

Tables and Figures

Table 1 consumer numbers by cluster, stratification levels, intervention step and timepoint

(a) Stream-one trial numbers of consumer participants in the three cross-sectional surveys who completed the Questionnaire about the Process of Recovery (QPR)

	Site information				QPR surveys		
Cluster	**Organisation**	**Strata**	**N**	**%**	**T0**	**T1**	**T2**
1	PMHS	A	66	7·0	23	29	14
2	PMHS	B	37	3·9	14	12	11
3	PMHS	C	66	7·0	21	24	21
4	PMHS	D	104	11·0	32	38	34
5	MHCSS - 1	E	52	5·5	16	15	21
6	MHCSS – 2	F	64	6·8	20	17	17
7	MHCSS - 1	G	56	5·9	19	25	12
8	PMHS	A	98	10·4	30	26	42
9	PMHS	B	44	4·7	17	15	12
10	PMHS	C	89	9·5	21	41	27
11	PMHS	D	71	7·5	20	26	25
12	MHCSS - 1	E	69	7·3	21	24	24
13	MHCSS – 2	F	52	5·5	21	17	14
14	MHCSS – 2	G	74	7·9	26	25	23
	Total, All sectors		***942***	***100***	***301***	***334***	***307***

Notes. Clusters were stratified by the team/service type and composition: i.e. seven different strata. Overall there were 575 (61·0%) consumer QPR surveys from Public Mental Health Services (PMHS) and 367 (39·0%) from Mental Health Community Support Services (MHCSS): 177 (18·8%) from MHCSS-1 and 190 (20·2%) from MHCSS-2.

(b) Stream-two trial numbers of consumer participants who participated in a study interview

	Site information				QPR surveys		
Cluster	**Organisation**	**Strata**	**N**	**%**	**T0**	**T1**	**T2**
1	PMHS	A	22	8·1	10	12	-
2	PMHS	B	15	5·5	10	5	-
3	PMHS	C	17	6·2	6	11	-
4	PMHS	D	24	8·8	14	10	-
5	MHCSS - 1	E	11	4·0	9	2	-
6	MHCSS – 2	F	19	7·0	11	8	-
7	MHCSS - 1	G	23	8·4	11	12	-
8	PMHS	A	26	9·5	-	13	13
9	PMHS	B	7	2·3	-	5	2
10	PMHS	C	29	10·6	-	12	17
11	PMHS	D	16	5·7	-	9	7
12	MHCSS - 1	E	23	8·4	-	9	14
13	MHCSS – 2	F	18	6·6	-	12	6
14	MHCSS – 2	G	23	8·4	-	9	14
	Total, All sectors		***273***	***100***	***71***	***129***	***73***

Notes. Clusters were stratified by the team/service type and composition: i.e. seven different strata. Overall there were 156 (57·1%) interviews with consumers from Public Mental Health Services (PMHS) and 117 (42·9%) from Mental Health Community Support Services (MHCSS): 57 (20·9%) from MHCSS-1 and 60 (22%) from MHCSS-2.

Key:
Control condition period
Intervention condition period

Table 2 Stream-one trial QPR numbers (%) by timepoint, gender, age group, step, intervention status and demographics

	Timepoint			
	T0	**T1**	**T2**	**Total**
Distribution in specialist care by Timepoint				
N	301	334	307	942
(%)	(32·0)	(35·5)	(32·6)	(100)
Distribution in specialist care by Timepoint and Gender[1]				
Female	174 (57·8)	192 (57·5)	178 (58·0)	544 (57·7)
Male	125 (41·5)	139 (41·6)	126 (41·0)	390 (41·4)
Not listed	2 (0·7)	3 (0·9)	3 (1·0)	8 (0·8)
Distribution in specialist care by Timepoint and Age group				
17-30 years	73 (24·3)	77 (23·1)	79 (25·7)	229 (24·3)
30-49 years	151 (50·2)	170 (50·9)	151 (49·2)	472 (50·1)
50 years and over	72 (23·9)	84 (25·1)	74 (24·1)	230 (24·4)
Distribution in specialist care by Timepoint and Step Group intervention				
Step Group 1	145 (48·2)	160 (47·9)	140 (45·6)	445 (47·2)
Step Group 2	156 (51·8)	174 (52·1)	167 (54·4)	497 (52·8)
Distribution in specialist care by Timepoint and Intervention status (Ix)				
No Ix	301 (100)	174 (52·1)	0 (0·0)	475 (50·4)
Yes Ix	0 (0)	160 (49·9)	307 (100)	467 (49·6)
Distribution in specialist care by Country of birth				
Australia	217 (72·1)	244 (73·1)	229 (74·6)	690 (73·2)
Other	83 (27·6)	87 (26·0)	73 (23·8)	243 (25·8)
Not listed	1 (0·4)	3 (0·9)	5 (1·6)	9 (1·0)
Distribution in specialist care by Year of arrival				
After 2000	17 (5·6)	23 (6·9)	19 (6·2)	59 (6·3)
Between 1981-2000	40 (13·3)	39 (11·7)	27 (8·8)	106 (11·3)
Before 1980	18 (6·0)	17 (5·1)	17 (5·5)	52 (5·2)
Not listed	8 (2·7)	8 (2·4)	10 (3·3)	26 (2·8)
Distribution in specialist care by Main language				
English	265 (88·0)	286 (85·6)	269 (87·6)	820 (87·0)
Other	23 (7·6)	26 (7·8)	23 (7·5)	72 (7·6)
Both English and Other	8 (2·7)	17 (5·1)	7 (2·3)	32 (3·4)
Not listed	5 (1·7)	5 (1·5)	8 (2·6)	18 (1·9)
Distribution in specialist care by Ethnicity (self-identified)				
Australian Non-Indigenous	121 (40·2)	177 (53·0)	162 (52·8)	460 (48·8)
Australian Indigenous	27 (9·0)	20 (6·0)	33 (10·7)	80 (8·5)
Other	120 (39·9)	126 (37·7)	97 (31·6)	343 (36·4)
Not listed	33 (11·0)	11 (3·3)	15 (4·9)	59 (6·3)
Other category (multiple responses could be listed)				
English, Irish, Walsh, Scottish	25 (8·3)	42 (12·6)	29 (9·4)	96 (10·2)
Italian	13 (4·3)	17 (5·1)	10 (3·3)	40 (4·2)
Greek	7 (2·3)	17 (5·1)	11(3·6)	35 (3·7)
New Zealander/Maori	11 (3·7)	10 (3·0)	12 (3·9)	33 (3·5)
Other (participant selected "other")	72 (23·9)	58 (17·4)	33 (10·7)	163 (17·3)
Censored[2]	108 (35·9)	91 (27·2)	76 (24·8)	275 (29·2)
Distribution in specialist care by Duration of mental health service use				
Mean number of years	4·0	4·5	4·0	4·2
Median number of years	1·0	1·0	1·0	1·0
Range (years)	0-35	0-35	0-35	0-35
No. of people with <1 year at site	129 (42·9)	125 (37·4)	135 (44·0)	389 (42·3)
Mean number of months for those with <1 year at site	3·3	3·2	3·2	3·2
Median number of months for those with <1 year at site	3	3	3	3

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.
[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other _____.
[2]Included 56 additional ethnic groups

Table 3 Stream-one QPR mixed model coefficients with fixed factors of sector, step-group, sex, age-group, time-point and intervention status, with clusters as random. Number of observations=942

		b	se	z	p value	ll	ul
sex							
	Female	-0·81	1·04	-0·79	·431	-2·84	1·21
Age group							
	2	-0·94	0·88	-1·07	·285	-2·65	0·78
	3	-3·4	0·91	-3·78	0·001***	-5·22	-1·66
Timepoint							
	T1	-3·22	1·02	-3·16	·002***	-5·22	-1·22
	T2	-4·22	1·50	-2·82	·005***	-7·15	-1·29
Intervention status							
	yes	3·76	1·31	2·87	·004***	1·20	6·33
Sector							
	2	-1·72	2·12	-0·81	·418	-5·87	2·43
Step group							
	2	0·15	2·08	0·07	·943	-3·93	4·22

*** p<0·01, ** p<0·05, * p<0·1

Table 4 Steam-one model-adjusted QPR means - derived from model in Table 3

		QPR mean	Std.Err.	[95%Conf.	Interval]
sex					
	1	54·85	1·36	52·18	57·51
	2	54·03	1·23	51·63	56·44
Age group					
	1	55·69	1·17	53·39	57·99
	2	54·75	1·25	52·31	57·20
	3	52·25	1·48	49·34	55·16
Timepoint					
	0	56·89	1·25	54·45	59·34
	1	53·67	1·43	50·86	56·48
	2	52·67	1·38	49·97	55·37
Intervention status					
	0	52·51	1·46	49·65	55·37
	1	56·27	1·23	53·87	58·67
Sector					
	1	55·05	1·75	51·62	58·47
	2	53·33	1·24	50·91	55·75
Step					
	1	54·29	1·42	51·52	57·07
	2	54·44	1·69	51·12	57·76

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0·5 – 6·8).

Table 5 Summary of outcomes in the streams 1 and 2 trials

Primary outcome:		**Control (n=475)**	**Intervention (n=467)**	**Adjusted diff in means (95%CI); p-value**	**Change in direction of favouring the intervention**
QPR Stream-one	Mean (sd)	53·6 (16·3); n=475	54·4 (16·2); n=467	3·72 [d] (0·51,6·92); 0·023	Yes
Secondary outcomes:		**Control (n=140)**	**Intervention (n=133)**		
QPR Stream-two	Mean (sd)	53·1 (14·8); n=138	54·0 (14·5); n=131	2·54[d] (-3·10,8·18); 0·38	
Warrick	Mean (sd)	41·4 (11·2); n=139	42·2 (11·1); n=133	2·39 [d] (-2·66,7·43); 0·35	Yes
INSPIRE S score	Mean (sd)	62·4 (22·3); n=128	62·2 (23·1); n=123	2·03 [d] (-6·72,10·78); 0·65	Yes
INSPIRE R score	Mean (sd)	72·0 (22·3); n=134	75·5 (20·1); n=129	3·29 [d] (-3·39,9·97); 0·34	
Other outcomes:		**Control (n=140)**	**Intervention (n=133)**		
GAF score	Mean (sd)	48·5 (14·7); n=140	51·4 (13·3); n=133	0·92 [d] (-6·15, 8·00); 0·80	Yes
SOFA score	Mean (sd)	49·8 (15·5); n=134	52·9 (14·3); n=132	0·57 [d] (-5·30, 6·45); 0·85	Yes
Client Satisfaction Questionnaire (CSQ)	Mean (sd)	23·3 (5·3); n=139	24·5 (5·5); n=130	1·21 [d] (-0·98, 3·41); 0·28	Yes
Mind Australia Satisfaction Survey (MASS)	Mean (sd)	8·0 (1·8); n=140	8·2 (1·8); n=132	0·02 [d] (-0·62, 0·67); 0·94	Yes
The Coercion Ladder, Community services	Median (IQR)	2·0 (1·5); n=139	2·0 (1·5); n=139	0·20 [II] (-1·12, 0·72); 0·67	Yes
Days out of role (full)	Median (IQR)	6·5 (0·0,15·0); n=138	6·0 (0·0,15·0); n=133	-1·37 (-5·34, 2·59); 0·50	No
Days out of role (partial)	Median (IQR)	6·0 (0·0,15·0); n=133	10·0 (2·0,15·0); n=129	0·12 (-4·56, 4·81); 0·96	
PNCQ (see Appendix 7)					Yes

Mean and standard deviation (sd) unless otherwise indicated. Also shown are the adjusted differences calculated from the multi-level mixed models (d linear or II Poisson regressions) adjusted for fixed effects of gender, age, timepoint and sector and clusters as random effects. Step group is an additional fixed effect in the stream-one regressions.

Figures

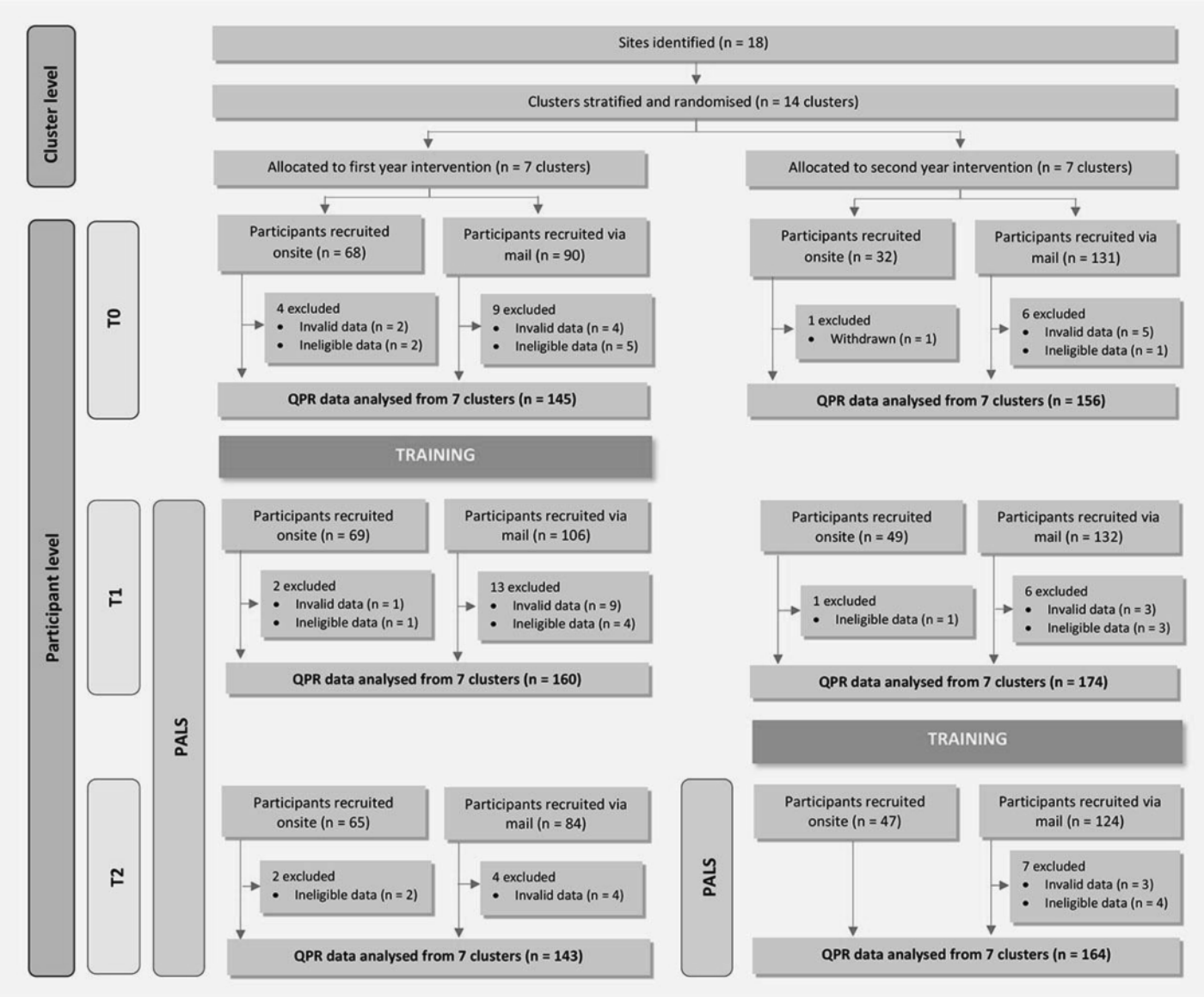


Figure 1 Consort chart for stream-one

Notes. PALS = PULSAR Active Learning Sessions. Invalid data refers to data-based issues in the form of missing data or invalid responses. Ineligible data refers to participant-based issues – that is, the person providing the data did not meet the eligibility criteria for the study.

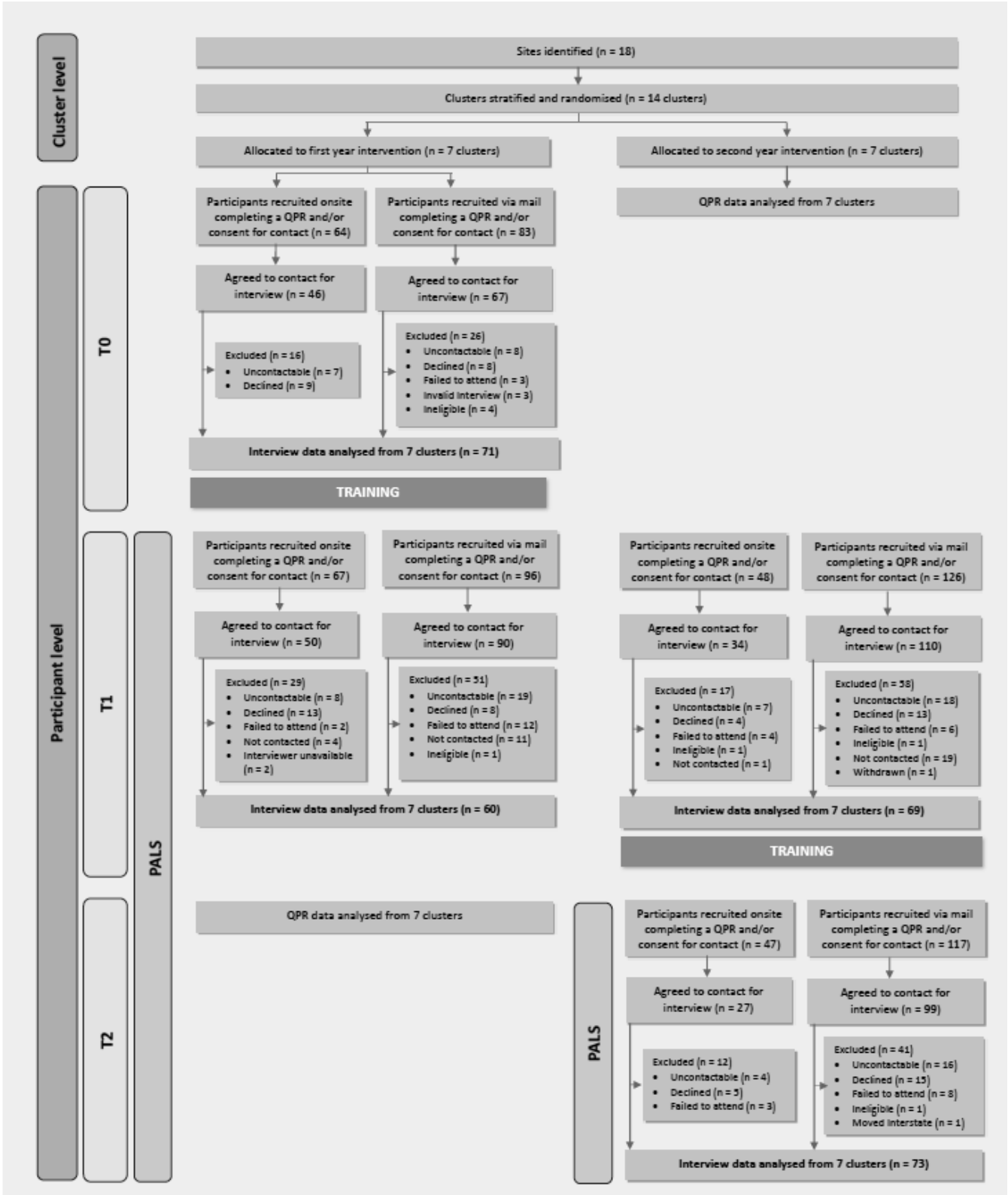


Figure 2 Consort chart for stream-two
Note. PALS = PULSAR Active Learning Sessions

Figure 3a: Model adjusted mean QPR scores over time: Public Mental Health Services (PMHS)

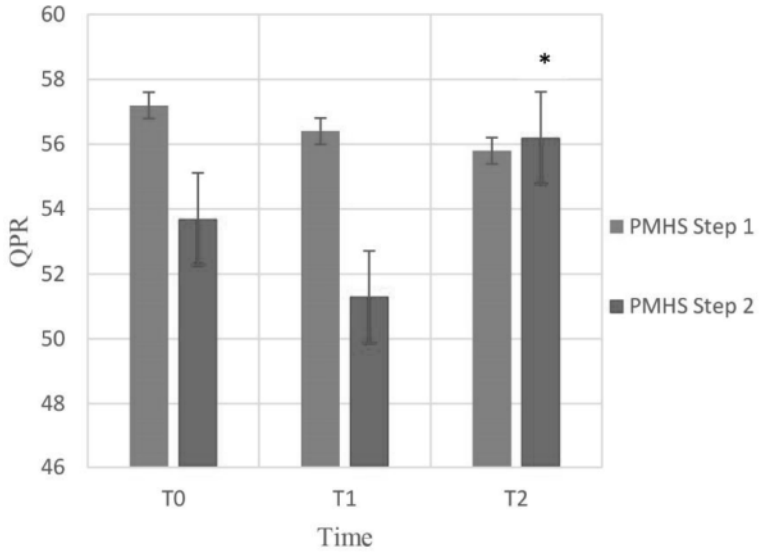


Figure 3b: Model adjusted mean QPR scores over time: Mental Health Community Support Services (MHCSS)

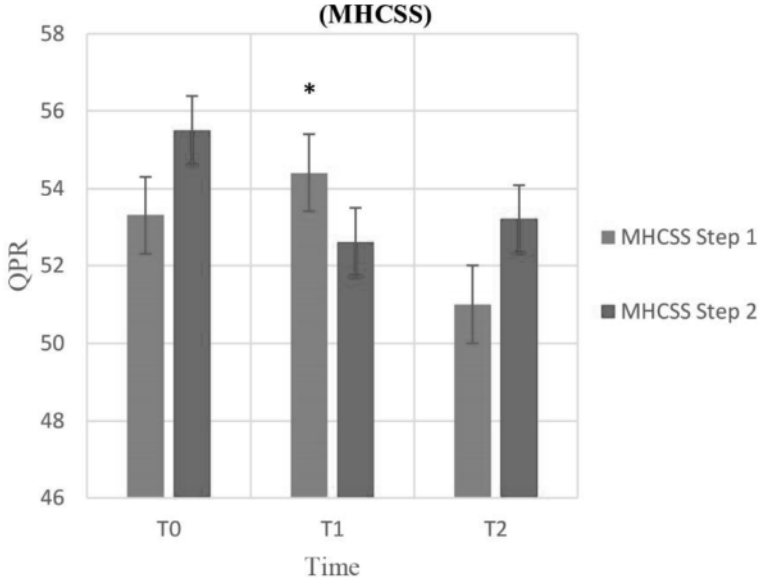


Figure 3 QPR scores by sector over time.

*Change $p < .01$ by pairwise comparison with previous time-point.
Note. Step-one group (blue) received intervention in year 1. Step-two group (red) received intervention in year 2.

References

1. Deegan P. Recovery as a journey of the heart. *Psychiatr Rehabil J* 1996; **19**(3): 91-7.

2. Slade M, Bird V, Le Boutillier C, et al. Development of the REFOCUS intervention to increase mental health team support for personal recovery. *Br J Psychiatry* 2015; **207**(6): 544-50.

3. Shepherd G, Boardman J, Slade M. Making recovery a reality. London: Sainsbury Centre for Mental Health 2008.

4. World Health Organization. Mental health action plan 2013 - 2020. Geneva: WHO Document Production Services; 2013.

5. Australian Government Department of Health. National mental health policy 2008. Canberra: Commonwealth of Australia 2009.

6. Slade M, Bird V, Clarke E, et al. Supporting recovery in patients with psychosis through care by community-based adult mental health teams (REFOCUS): a multisite, cluster, randomised, controlled trial. *Lancet Psychiatry* 2015; **2**(6): 503-14.

7. Salkeld R, Wagstaff C, Tew J. Toward a new way of relating: an evaluation of recovery training delivered jointly to service users and staff. *J Ment Health* 2013; **22**(2): 165-73.

8. Deane FP, Andresen R, Crowe TP, Oades LG, Ciarrochi J, Williams V. A comparison of two coaching approaches to enhance implementation of a recovery-oriented service model. *Adm Policy Ment Health* 2014; **41**(5): 660-7.

9. Uppal S, Oades LG, Crowe TP, Deane FP. Barriers to transfer of collaborative recovery training into Australian mental health services: implications for the development of evidence-based services. *J Eval Clin Pract* 2010; **16**(3): 451-5.

10. Lodge AC, Kaufman L, Stevens Manser S. Barriers to Implementing Person-Centered Recovery Planning in Public Mental Health Organizations in Texas: Results from Nine Focus Groups. *Adm Policy Ment Health* 2017; **44**(3): 413-29.

11. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs. The Four Levels. Third ed. San Franciso, California: Berrett-Koehler Publishers; 2006.

12. Slade M, Bird V, Le Boutillier C, Williams J, McCrone P, Leamy M. REFOCUS Trial: protocol for a cluster randomised controlled trial of a pro-recovery intervention within community based mental health teams. *BMC Psychiatry* 2011; **11**(1): 185.

13. Ajzen I. The theory of planned behaviour: reactions and reflections. *Psychol Health* 2011; **26**: 1113–27.

14. Le Boutillier C, Chevalier A, Lawrence V, et al. Staff understanding of recovery-orientated mental health practice: a systematic review and narrative synthesis. *Implement Sci* 2015; **10**(1): 1-14.

15. Neil ST, Kilbride M, Pitt L, et al. The questionnaire about the process of recovery (QPR): A measurement tool developed in collaboration with service users. *Psychosis* 2009; **1**(2): 145-55.

16. Chow SC, Chang M. Adaptive design methods in clinical trials - a review. *Orphanet J Rare Dis* 2008; **3**: 11.

17. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. *BMJ Quali Saf* 2014; **23**: 290-8.

18. Bird V, Leamy M, Le Boutillier C, Williams J, Slade M. REFOCUS: Promoting Recovery in Mental Health Services. 2nd ed. London: Institute of Psychiatry, King's College London; 2014.

19. Shawyer F, Enticott JC, Brophy L, et al. The PULSAR Specialist Care protocol: a stepped-wedge cluster randomized control trial a training intervention for community mental health teams in recovery-oriented practice. *BMC Psychiatry* 2017; **17**: 172.

20. PULSAR-REFOCUS team. PULSAR Manual: Recovery-promoting Relationships and Working Practices for Specialist and Community Mental Health Services (or Secondary Care Staff). 2nd ed. Melbourne: Monash University; 2016.

21. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions an introduction to the new Medical Research Council guidance. In: Kelly MP, Killoran A, eds. Evidence–Based Public Health: Effectiveness and Efficiency. Oxford: Oxford University Press; 2010: 185-203.

22. Australian Government. National Standards for Mental Health Services 2010. Australia: Commonwealth of Australia; 2010.

23. Williams J, Leamy M, Pesola F, Bird V, Le Boutillier C, Slade M. Psychometric evaluation of the Questionnaire about the Process of Recovery (QPR). *Br J Psychiatry* 2015; **207**(6): 551-5.

24. Williams J, Leamy M, Bird V, et al. Development and evaluation of the INSPIRE measure of staff support for personal recovery. *Soc Psychiatry Psychiatr Epidemiol* 2015; **50**(5): 777-86.

25. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). *Soc Psychiatry Psychiatr Epidemiol* 2000; **35**(9): 427-35.

26. Attkisson CC, Zwick R. The Client Satisfaction Questionnaire: psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982; **5**(3): 233-7.

27. Wray M, Brophy L, Loh M. The Prevention and Recovery Care Exit Survey: Mind Australia; 2013.

28. Høyer G, Kjellin L, Engberg M, et al. Paternalism and autonomy: A presentation of a Nordic study on the use of coercion in the mental health care system. *Int J Law Psychiatry* 2002; **25**(2): 93-108.

29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). Washington, DC, USA: American Psychiatric Association;; 2000.

30. Hemming K, Girling A. A menu-driven facility for power and detectable-difference calculations in stepped-wedge cluster-randomized trials. *Stata J* 2014; **14**(2): 363-80.

31. Statacorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.

32. McNeish DM, Stapleton LM. The Effect of Small Sample Size on Two-Level Model Estimates: A Review and Illustration. *Educ Psychol Rev* 2016; **28**(2): 295-314.

33. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007; **28**(2): 182-91.

34. Hemming K, Taljaard M, Frobes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. *Trials* 2017; **18**(101).

35. Patsopoulos N. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosc* 2011; **13**(2): 217-24.

36. Selya AS, Rose JS, Dierker LC, Hedeker D, Mermelstein RJ. A practical guide to calculating Cohen's f(2), a measure of local effect size, from PROC MIXED. *Front Psychol* 2012.

37. Brysbaert M, Stevens M. Power analysis and effect size in mixed effects models: A tutorial. *J Cogn* 2018; **1**(1).

38. Sharkey K, Savulescu J, Aranda S, Schofield P. Clinician gate-keeping in clinical research is not ethically defensible: an analysis. *J Med Ethics* 2010; **36**(6): 363-6.

39. Stuber J, Rocha A, Christian A, Johnson D. Predictors of Recovery-Oriented Competencies Among Mental Health Professionals in One Community Mental Health System. *Community Ment Hlt J* 2014; **50**(8): 909-14.

40. Le Boutillier C, Leamy M, Bird VJ, Davidson L, Williams J, Slade M. What does recovery mean in practice? A qualitative analysis of international recovery-oriented practice guidance. *Psychiatr Serv* 2011; **62**(12): 1470-6.

Supplementary materials

Appendix 1

Key changes in the setting of the project through the time-period for observations

The key period in which this intervention and associated observations across multiple healthcare sectors in Victoria occurred from late 2014 to early 2017 was a time of considerable external change, collective stresses and challenge to involved organisations and their staff. This context is likely to have had some negative influences on implementation of recovery orientated practice.

The funding environment for public health services in Victoria under the Liberal administration 2010 to 2015 received significant criticism as negative, with funding levels not keeping up with inflation, or where they did, being associated with substantial additional commitments so not representing real increases.[1] While the Labor administration that followed has been better reviewed for its support of healthcare,[2] there was limited time for the actions of this new administration to flow through into changes in work context in the timespan of this project.

As well as the general problem of under resourcing, three intersecting areas of change impacted on this project and the research undertaken with specialist mental health services.

Changes to the MHCSS:

In 2015, the then Victorian State Government introduced a major reform of mental health community support services (MHCSS), which presented substantial challenges for organisations involved in this study.[3] In 2014, Kim Koop, the then CEO of the peak body for mental health community support services (MHCSS) in Victoria, anticipated that MHCSS were about to experience "an extended period of uncertainty" (p.10)[4] as a result of new service types and new contracts being implemented to begin to reshape service delivery that would also herald significant changes in the service provider landscape.[4] This process was seen as necessary to deal with a long-standing need for reform and also to assist the sector to prepare for the much larger transition process of having all MHCSS funding rolling into the National Disability Insurance Scheme (NDIS) by 2019. Later referred to as "recommissioning", Kate Silburn wrote, after interviewing senior personnel, that:

> In the case of recommissioning MHCSS and AOD (Alcohol and other drug) sectors in Victoria there would appear to have been very high costs, potentially avoidable difficulties and as yet undetermined benefits. One of strongest themes emerging from the interviews was that while there had been widespread enthusiasm for reform in both sectors, many of its proponents had become disillusioned and despairing as a consequence of the processes for both reform design and reform implementation (p.39).[3]

The recommissioning involved the establishment of new catchment areas, a centralised intake and assessment process, and a reduction in the number of providers. Many small services lost their funding altogether and larger services had both losses and gains that required considerable adjustment. For example some services were allocated to other providers and staff and clients needed to transition from one provider to another. This resulted in a long and difficult transition period. The costs to the sector were considerable. Agencies, including those in this study, had to make staff redundant and also explain to consumers why they were no longer going to provide services to them. Other problems identified by Silburn[3] are:

- Poor planning, lack of guidelines and information and limited systems alliance (p.19)
- Not enough time and too many things happening at once (p.21)
- Lack of communication with consumers and other types of service providers (p.21)
- Lack of a well-planned process for 'transitioning' clients (p.21)

Silburn (2014) also found that the recommissioning process had undermined:

- collaboration, partnerships and joint models of care (p.24);
- models of care for clients from disadvantaged or vulnerable communities (p.25);
- comprehensive models of care delivered by single agencies

There were concerns that the central intake system became a barrier for clients to access services because it was often multi-stepped and difficult to negotiate. For example:

> Interviewees argued that while clients had previously been able to walk into their agency, make an appointment and get an assessment within a short time period, they now had to be directed to call the central intake provider and may have to wait several weeks for an assessment (p.29)[3]

After recommissioning Silburn[3] describes how:

> MHCSS sector clients are categorised into three tiers, consistent with the proposed categorisation for the NDIS. To be eligible for a service clients have to have a permanent disability associated with a mental illness. Once clients are deemed eligible they are then categorised based on the severity of the disability and/or their current needs. This means that clients with high levels of disability, but who are otherwise stable/doing well (and therefore might have low levels of need) might get the same level of priority as someone who has a lower level of disability and a high level of current need. One interviewee noted that this also means that clients with either no permanent disability or with an uncertain diagnosis, but with high needs are likely to miss out on services in the new system. People in this group can include people experiencing their first psychosis or life circumstances like deterioration of their support networks, who with early intervention may not become

dependent on the MHCSS system. The system has therefore lost significant capacity for prevention for this group of clients. (p.32).

Changes as a result of the new Mental Health Act:

In 2014, the Victorian State government introduced a new Mental Health Act.[5] Although the Act is encouraging of a 'recovery' orientation, this is not explicitly defined in the Act, and was introduced along with its further training requirements in the context of this budget-limited environment. The introduction of the new Mental Health Act in 2014 was the first such change in nearly three decades. It provided an entirely new legal framework for the delivery of mental health services (MHS) and carried with it considerable training demands on clinicians in the public MHS involved. This led to some delay in the ability to deliver the PULSAR training since it was not practicable to release staff in those services so quickly following this other major training impost. The MHA was designed to support recovery and introduce a range of new mechanisms to improve patients or consumers involvement in decision making including Advance Statements and Nominated Persons. The immediate impact of the legislative changes was to produce an uncertain legal environment in which the voluntary PULSAR training modules competed with compulsory training on the Act for staff time and attention. Also this compulsory training tended not to focus on the relevant recovery and supported decision making reforms but rather on the changes to compulsory admission criteria and treatment orders, restrictive interventions and ECT.

Changes at the Clinical Services:

Some quantitative summary data gives an indication of the trends in activity through time across the PMHS, based on regularly collected data and available reports which are not available in as standardised a way for the MHCSS. The last day of the year snapshot of all PMHS case managed clients rose from 2349 in 2014 to 2462 in 2016, an increment of 5%. By way of indicators on demand factors for the whole service, emergency department presentations increased from 8803 in 2014 to 10004 in 2016 (+14%) and inpatient length of stay decreased on average from 12·8 days to 11·3 days (-4%) as total in-patient separations increased from 3102 to 3633 (+17%). Average length of stay in community services from opening to closing of administrative cases increased by 31% (2014: 157 days, 2015: 170 days, 2016: 205 days).
From an observational and more qualitative perspective, Monash Health Service (MH) could be described as a hyper-complex environment [6] and it was particularly so during the time this study was being conducted. In 2015, among significant changes within MH through the course of the project, it is publicly available information that the staff employed as MH Medical Program Director and Executive Director left in April and May of that year respectively. Long et al[7] describe, in reporting on a project that was also occurring during this time, how "MHS underwent a major restructuring after a significant number of senior staff left the service" (p.2).
Long, McDermott and Meadows (as yet unpublished PhD research[8]) describe, via semi structured interviews carried out with the MHS senior leadership group, the amount of change and challenges occurring in the service between 2013 and 2017. While their investigation was a different project to PULSAR it occurred in a similar timeframe and the findings are very relevant to describing the context. Long et al's participants used critical reflection to identify meaningful events in the services during this time.[8] Twenty-three critical incidents were identified. These included changes in government policy, adjustments in funding and staff turnover. Hence staff in the service were persistently having to deal with change and also the loss of some programs, creating an atmosphere of uncertainty.

Appendix 2

Training quality assurance and adaptations made to the training in course of the project

Quality assurance measures employed during training delivery included a day-long workshop attended by all trainers to introduce the training schedule, content and process including demonstrations and role-play of key exercises. A detailed schedule guided delivery working through the key elements of content along with use of standardised training materials including a range of consistently employed audio-visual aids. Discussions with a CI early in the training schedules followed each day session to review any departures from intended process- as confidence grew with the training these were replaced with accessibility of a CI to discuss any problems following the sessions.
The first intervention round for clinical services was developed as a two-day session, with the community services training planned as a separate two-day session in the same week. In addition to two consumer trainers employed by the project, trainers were sourced from clinical services for the clinical sessions and the community sector for the community sessions. This was anticipated as enabling the inclusion of specialist skills and experience in the delivery of training.
Training in the second round was subject to further modifications based on analyses of evaluations of the first round of training by both participants and trainers. The delivery of the intervention was modified to account for previously unknown restrictions on the ability of services to release staff for two days of training. Based on feedback from services, it was identified that attending two days of training for some teams was difficult. This was either due to the workload of the teams (specifically CAT teams) or the recent undertaking of organisational wide recovery training. In response to this the training was re-designed so that all material is covered in the first day of training, with more in-depth exploration and practice of the knowledge and skill on day two.

Feedback from the first round of training both through the structured feedback following training and from qualitative work led us to make several other modifications:
Training was restructured to allow half of the two days of training to be combined between the MHCSS sector and MH Staff. Feedback highlighted how the consumer role in leading training could be experienced as very challenging for some participants particularly if the consumer was experienced as critical of staff. Of course, being open to hearing criticisms from consumers about mental health care is a critical part of any transition to recovery-oriented practice so the training team worked very hard at considering this feedback in subsequent rounds. A key learning was that the introduction of the REACH coaching process needed to be deeply experiential. In particular, the training team formed the view that a critical element of the delivery was that the co-trainers as consumers and clinicians of other workers needed to embody the coaching principles in a fully authentic way. In alignment with a PDSA approach we took this on board as much as we could and adjusted the interactive style of the trainers for the second round. Specific focuses coaching based material was added with involvement of an additional trainer providing this particular perspective. Additions to the PULSAR Manual included sections providing information on Advance statements, Nominated persons and Risk and Recovery with additional references. Additional material was provided in Appendix 2 and the title changed from "Additional resources for understanding values" to "Additional resources for consumers' experiences". Additional web resources were added to Appendix 7 and the title changed from "Example of a relapse symptoms checklist" to "Care plans, and example of a relapse symptom checklist and other resources".

Appendix 3

Adverse events

At the commencement of the trial the forms of possible adverse events we anticipated included: 1) risk of distress by a participant during an interview; 2) issues related to disclosure of potential harm to self or others 3) risk of harm to staff. We developed an ethics protocol outlining the prevention and management of these risks which was approved by the governing HREC and our participant information and consent form for the face-to-face interviews outlined the potential risk of distress and what to do should it occur. Participants who were invited to complete the survey or undertake and interview were provided with written contact details for complaints, which was the manager of the governing HREC. During the course of the project there were four complaints reported to the HREC. Three complaints related to QPR mailout (privacy concerns; receiving a letter but not a client of participating services; receiving a letter to a consumer who had died) which led to changes in procedures under direction of HREC as appropriate. One complaint related to the management of interview distress which led to updates to the staff training protocol and counselling provided to the staff member concerned.
No complaints were received that related to the REFOCUS-PULSAR intervention.
In addition to these complaints reported to HREC, during the trial one participant expressed suicidal ideation in a note attached to the return of a consent form. This was followed up by staff as per our ethics protocol to ensure their safety. The suicidal ideation was related to chronic psychiatric and medical symptoms and family conflict, not to participation in the project. In line with CATT advice, this issue was ultimately passed onto the police who took the participant to hospital, as the participant could not guarantee her safety.
Beyond the complaints process, we did not systematically collect any other adverse event information from consumers (such as deaths, hospitalisations etc.) because the intervention was a training intervention for staff rather than a clinical intervention for consumers so attribution of adverse events from clinical care in the context of the study, of which care guided by PULSAR-REFOCUS principles was but a part would not have been clear.

Appendix 4

Figure: Detailed CONSORT chart for stream-one

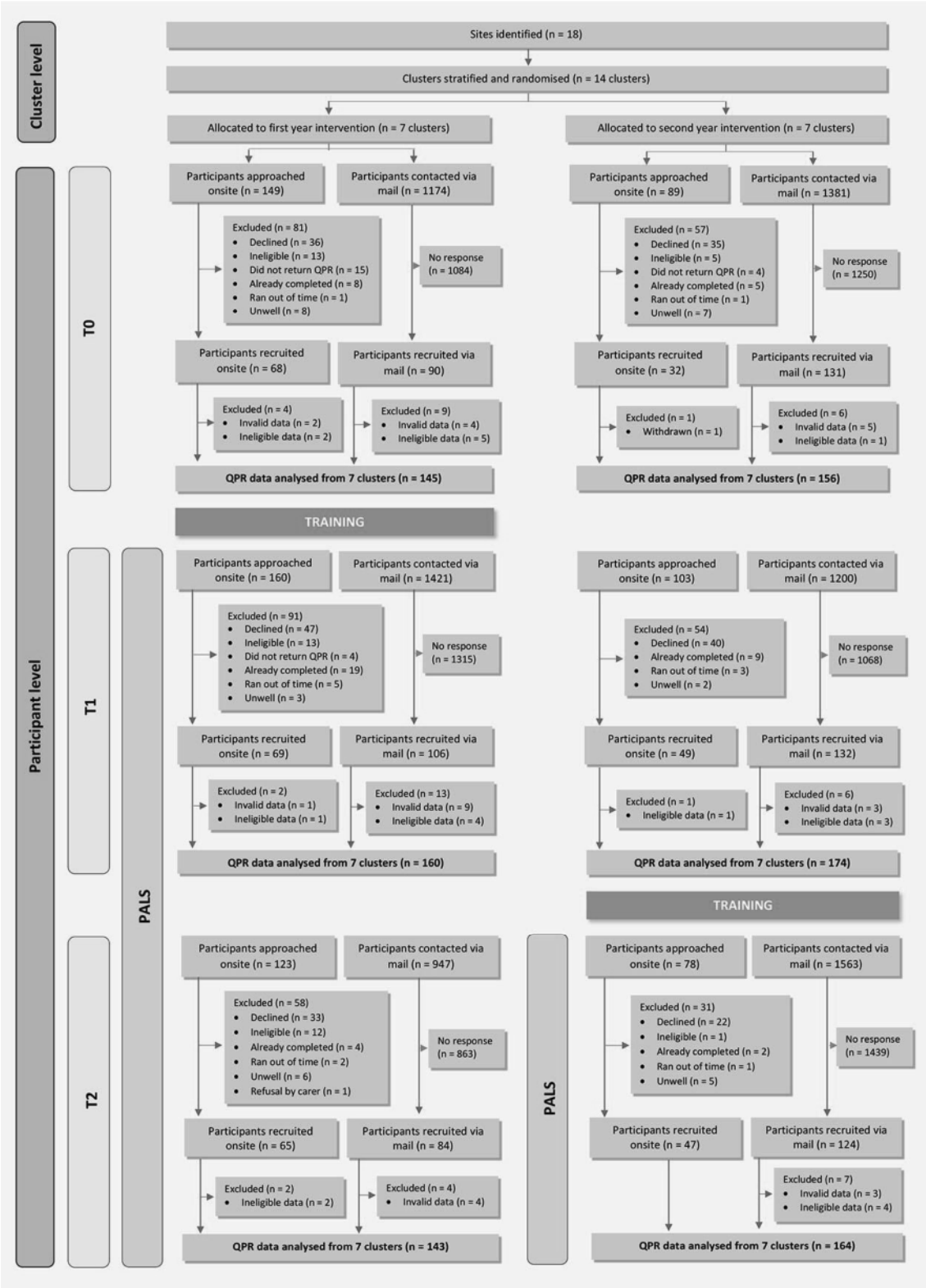


Figure S1: Detailed CONSORT chart stream-one

Notes. PALS = PULSAR Active Learning Sessions. Invalid data refers to data-based issues in the form of missing data or invalid responses. Ineligible data refers to participant-based issues – that is, the person providing the data did not meet the eligibility criteria for the study.

Appendix 5

Details on profile of participants

Table 5.1 Stream-one trial QPR numbers (%) by Intervention Status, gender, age group, step, intervention status and demographics.

There were no significant differences between the control and intervention groups.

	Control	Intervention	Total
Distribution in specialist care			
N	475	467	942
(%)	(50·4·0)	(49·6)	(100)
Distribution in specialist care by Gender[1]			
Female	268 (56·4)	276 (59·1)	544 (57·7)
Male	203 (42·7)	187 (40·0)	390 (41·4)
Not listed	4 (0·8)	4 (0·9)	8 (0·8)
Distribution in specialist care by Age group			
17-30 years	104 (22·2)	125 (26·1)	229 (24·3)
30-49 years	243 (51·8)	229 (49·6)	472 (50·1)
50 years and over	122 (26·0)	108 (23·4)	230 (24·4)
Distribution in specialist care by Step Group intervention			
Step Group 1	145 (30·5)	300 (64·2)	445 (47·2)
Step Group 2	330 (69·5)	167 (35·8)	497 (52·8)
Distribution in specialist care by Intervention status (Ix)			
No Ix	475 (100)	0 (0·0)	475 (50·4)
Yes Ix	0 (0)	467 (100)	467 (49·6)
Distribution in specialist care by Country of birth			
Australia	345 (72·6)	345 (73·9)	690 (73·2)
Other[2]	125 (26·3)	118 (25·3)	243 (25·8)
Not listed	5 (1·1)	4 (0·9)	9 (1·0)
Distribution in specialist care by Year of arrival			
After 2000	27 (45·8)	32 (54·2)	59 (6·3)
Between 1981-2000	62 (58·5)	44 (41·5)	106 (11·3)
Before 1980	27 (54·0)	25 (46·0)	52 (5·2)
Not listed	12 (46·2)	14 (53·8)	26 (2·8)
Distribution in specialist care by Main language			
English	422 (88·8)	398 (85·2)	820 (87·0)
Other	32 (6·7)	40 (8·6)	72 (7·6)
Both English and Other	14 (2·9)	18 (3·9)	32 (3·4)
Not listed	7 (1·5)	11 (2·4)	18 (1·9)
Distribution in specialist care by Ethnicity (self-identified)			
Australian Non-Indigenous	223 (46·9)	237 (64·6)	460 (48·8)
Australian Indigenous	35 (7·4)	45 (12·3)	80 (8·5)
Other	176 (37·1)	167 (45·5)	343 (36·4)
Not listed	41 (8·6)	18 (4·9)	59 (6·3)
Distribution in specialist care by Duration of mental health service use			
Mean number of years	4·3	4·0	4·2
Median number of years	1·0	1·0	1·0
Range (years)	0-35	0-35	0-35
No. of people with <1 year at site	190 (40·0)	199 (42·6)	389 (42·3)
Mean number of months for those with <1 year at site	3·1	3·3	3·2
Median number of months for those with <1 year at site	3	3	3

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other _____

[2]Included 60 additional ethnic groups

Table 5.2 Stream-two trial QPR numbers (%) by timepoint, gender, age group, step, intervention status and demographics

	Timepoint			
	T0	T1	T2	Total
Distribution in specialist care by Timepoint				
N	71	129	73	273
(%)	(26·0)	(47·3)	(26·7)	(100)
Distribution in specialist care by Timepoint and Gender[1]				
Female	46 (64·8)	63 (48·8)	42 (57·5)	151 (55·3)
Male	25 (35·2)	64 (49·6)	31 (42·5)	120 (44)
Other	0 (0·0)	2 (1·6)	0 (0·0)	2 (0·7)
Distribution in specialist care by Timepoint and Age group				
17-30 years	26 (36·6)	23 (17·8)	13 (17·8)	62 (22·7)
30-49 years	30 (42·3)	67 (51·9)	35 (48)	132 (48·4)
50 years and over	15 (21·1)	38 (29·5)	24 (32·9)	77 (28·2)
Not listed	0 (0·0)	1 (0·8)	1 (1·4)	2 (0·7)
Distribution in specialist care by Timepoint and Step Group intervention				
Step Group 1	71 (100)	60 (46·5)	0 (0·0)	131 (48)

Step Group 2	0 (0·0)	69 (53·5)	73 (100)	142 (52)
Distribution in specialist care by Timepoint and Intervention status (Ix)				
No Ix	71 (100)	69 (53·5)	0 (0·0)	140 (51·3)
Yes Ix	0 (0·0)	60 (46·5)	73 (100)	133 (48·7)
Distribution in specialist care by Country of birth				
Australia	49 (69·0)	94 (72·9)	53 (72·6)	196 (71·8)
Other	22 (31·0)	35 (27·1)	20 (27·4)	77 (28·2)
Year of arrival in Australia				
After 2000	2 (9·1)	11 (18·3)	2 (10·0)	15 (19·5)
Between 1981-2000	12 (54·5)	14 (23·3)	10 (50·0)	36 (46·8)
Before 1980	6 (27·3)	10 (16·7)	7 (35·0)	23 (29·9)
Not listed	2 (9·1)	25 (41·7)	1 (5·0)	28 (36·4)
Distribution in specialist care by Main language spoken at home				
English	60 (84·5)	115 (89·2)	66 (90·4)	241 (88·3)
Other	11 (15·5)	14 (10·9)	7 (9·6)	32 (11·7)
Distribution in specialist care by Ethnicity (self-identified)				
Australian Non-Indigenous	38 (53·5)	74 (57·4)	43 (58·9)	155 (56·8)
Australian Indigenous	2 (2·8)	2 (1·6)	2 (2·7)	6 (2·2)
Other	20 (28·2)	50 (38·8)	19 (26)	89 (32·6)
Not listed	11 (15·5)	3 (2·3)	9 (12·3)	23 (8·4)
Other category (multiple responses could be listed)				
British (English, Irish, Walsh, Scottish)	1 (5·0)	17 (34·0)	5 (26·3)	23 (25·8)
European (Italian, Greek, Bosnian, Dutch, German)	7 (35·0)	20 (40·0)	5 (26·3)	32 (36·0)
New Zealander/Maori	2 (10·0)	3 (6·0)	4 (21·1)	9 (10·1)
Middle Eastern (Afghan)	2 (10·0)	0 (0·0)	0 (0·0)	2 (2·2)
South East Asian (Burmese, Chinese, Indian, Cambodian, Sri Lankan, Vietnamese)	8 (16·0)	5 (10·0)	5 (26·3)	18 (20·2)
Other (participant selected "other")	0 (0·0)	5 (10·0)	0 (0·0)	5 (5·6)
Distribution in specialist care by Duration of mental health service use				
Mean number of years	11·0	13·2	13·1	12·6
Median number of years	9·0	11·0	11·0	10·0
Range (years)	1-40	1-33	1-40	1-40
No. of people with <1 year at site	0	0	0	0
Duration of current service use				
Mean number of years	4·6	5·8	7·2	5·8
Median number of years	3·0	3·0	4·5	3·0
Range (years)	0-23	1-22	1-32	0-32
No. of people with <1 year at site	1	0	0	1
Median number of months for those with <1 year at site	3	-	-	3
Distribution in specialist care by Marital status				
Single	48 (67·6)	67 (51·9)	33 (45·2)	148 (54·2)
Married	6 (8·5)	21 (16·3)	16 (21·9)	43 (15·8)
DeFacto	3 (4·2)	8 (6·2)	1 (1·4)	12 (4·4)
Separated	4 (5·6)	15 (11·6)	10 (13·7)	29 (10·6)
Divorced	10 (14·1)	14 (10·9)	10 (13·7)	34 (46·6)
Widowed	0 (0·0)	1 (0·8)	2 (2·7)	3 (4·1)
Other	0 (0·0)	3 (2·3)	1 (1·4)	4 (5·5)
Distribution in specialist care by Child status				
Yes	33 (46·5)	64 (49·6)	41 (56·2)	138 (50·5)
No	38 (53·5)	65 (50·4)	32 (43·8)	135 (49·5)
Number of children living at home				
0	16 (48·5)	32 (50·0)	16 (39·0)	64 (46·4)
1	11 (33·3)	17 (26·6)	12 (29·3)	40 (29·0)
2	4 (12·2)	10 (15·6)	9 (22·0)	23 (16·7)
3	0 (0·0)	1 (1·6)	3 (7·3)	4 (2·9)
4	1 (3·0)	1 (1·6)	0 (0·0)	2 (1·4)
5-6	1 (3·0)	1 (1·6)	0 (0·0)	2 (1·4)
Not listed	0 (0·0)	2 (3·1)	1 (2·4)	3 (2·2)
Distribution in specialist care by Living situation (multiple responses could be selected)				
Living with parents	16 (22·5)	16 (12·4)	15 (20·5)	47 (17·2)
Living with siblings	4 (5·6)	8 (6·2)	4 (5·5)	16 (5·9)
Living with a partner	7 (9·9)	31 (24·0)	15 (20·5)	53 (19·4)
Living with children	15 (21·1)	26 (20·2)	19 (26·0)	60 (22·0)
Living with friends	4 (5·6)	7 (5·4)	2 (2·7)	13 (4·8)
Living in shared accommodation	5 (7·0)	14 (10·9)	5 (6·8)	24 (8·8)
Living in crisis accommodation	3 (4·2)	3 (2·3)	1 (1·4)	7 (2·6)
Living in support housing	8 (11·3)	11 (8·5)	9 (12·3)	28 (10·3)
Living alone	23 (32·4)	30 (23·3)	20 (27·4)	73 (26·7)
Homeless	3 (4·2)	3 (2·3)	0 (0·0)	6 (2·2)
Other	4 (5·6)	10 (7·8)	1 (1·4)	15 (5·5)
Distribution in specialist care by Education level				
None	0 (0·0)	0 (0·0)	1 (1·4)	1 (0·4)
Primary school	2 (2·8)	4 (3·1)	0 (0·0)	6 (2·2)
Secondary school (≤ yr 10)	25 (35·2)	39 (27·9)	21 (28·8)	85 (31·1)
Secondary school (yr 11)	12 (16·9)	30 (23·3)	14 (19·2)	56 (20·5)
Secondary school (yr 12)	27 (38·0)	56 (43·4)	37 (50·7)	120 (44·0)
Not listed	5 (7·0)	0 (0·0)	0 (0·0)	5 (1·8)

Distribution in specialist care by Highest qualification				
Certificate I	2 (2·8)	1 (0·8)	1 (1·4)	4 (1·5)
Certificate II	3 (4·2)	6 (4·7)	4 (5·5)	13 (4·8)
Certificate III	15 (21·1)	19 (14·7)	13 (17·8)	47 (17·2)
Certificate IV	7 (9·9)	15 (11·6)	10 (13·7)	32 (11·7)
Diploma	9 (12·7)	11 (8·5)	10 (13·7)	30 (11·0)
Advanced Diploma	1 (1·4)	1 (0·8)	1 (1·4)	3 (1·1)
Associate Degree	0 (0·0)	1 (0·8)	0 (0·0)	1 (0·4)
Bachelor Degree	1 (1·4)	13 (10·1)	7 (9·6)	21 (7·7)
Bachelor Honours Degree	0 (0·0)	3 (2·3)	1 (1·4)	4 (1·5)
Graduate Diploma	0 (0·0)	2 (1·6)	2 (2·7)	4 (1·5)
Masters (research)	0 (0·0)	1 (0·8)	0 (0·0)	1 (0·4)
Masters (coursework)	1 (1·4)	3 (2·3)	1 (1·4)	5 (1·8)
Doctoral	1 (1·4)	1 (0·8)	0 (0·0)	2 (0·7)
Other	3 (2·8)	11 (8·5)	4 (5·5)	18 (6·6)
Not listed	28 (39·4)	41 (31·8)	19 (26·0)	85 (31·1)

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other.

Table 5.3 Stream-two trial QPR numbers (%) by Intervention Status, gender, age group, step, intervention status and demographics.

There were no significant differences between the control and intervention groups.

	Control	Intervention	Total
Distribution in specialist care			
N	140	133	273
(%)	(51·3)	(48·7)	(100)
Distribution in specialist care by Gender[1]			
Female	80 (57·1)	71 (53·4)	151 (55·3)
Male	59 (42·1)	61 (45·9)	120 (44)
Not listed	1 (0·7)	1 (0·8)	2 (0·7)
Distribution in specialist care by Age group			
17-30 years	35 (25·0)	27 (20·3)	62 (22·7)
30-49 years	65 (46·4)	67 (50·4)	132 (48·4)
50 years and over	39 (27·9)	38 (28·6)	77 (28·2)
Not listed	1 (0·7)	1 (0·8)	2 (0·7)
Distribution in specialist care by Step Group intervention			
Step Group 1	71 (50·7)	60 (45·1)	131 (48·0)
Step Group 2	69 (49·3)	73 (54·9)	142 (52·0)
Distribution in specialist care by Intervention status (Ix)			
No Ix	140 (100)	0 (0·0)	140 (50·4)
Yes Ix	0 (0·0)	133 (100)	133 (49·6)
Distribution in specialist care by Country of birth			
Australia	98 (70·0)	98 (73·7)	196 (71·8)
Other[2]	42 (30·0)	35 (26·3)	77 (28·2)
Year of arrival in Australia			
After 2000	8 (19·0)	7 (20·0)	15 (19·5)
Between 1981-2000	20 (47·6)	16 (45·7)	36 (46·8)
Before 1980	12 (28·6)	11 (31·4)	23 (29·9)
Not listed	2 (4·8)	1 (2·9)	3 (3·9)
Distribution in specialist care by Main language			
English	124 (88·6)	117 (88·0)	241 (88·3)
Other	16 (11·4)	16 (12·0)	32 (11·7)
Distribution in specialist care by Ethnicity (self-identified)			
Australian Indigenous	2 (1·4)	4 (3·0)	6 (2·2)
Australian Non-Indigenous	77 (55·0)	78 (58·6)	155 (56·8)
Other	54 (38·6)	35 (26·3)	89 (32·6)
Not listed	7 (5·0)	16 (12·0)	23 (8·4)
Other category (multiple responses could be listed)			
British (English, Irish, Walsh, Scottish)	13 (24·1)	10 (28·6)	23 (25·8)
European (Italian, Greek, Bosnian, Dutch, German)	22 (40·7)	10 (28·6)	32 (36·0)
New Zealander/Maori	4 (7·4)	5 (14·3)	9 (10·1)
Middle Eastern (Afghan)	2 (3·7)	0 (0·0)	2 (2·2)
South East Asian (Burmese, Chinese, Indian, Cambodian, Sri Lankan, Vietnamese)	13 (24·1)	10 (28·6)	23 (25·8)
Other (participant selected "other")	0 (0·0)	0 (0·0)	0 (0·0)
Distribution in specialist care by Duration of mental health service use			
Mean number of years	13·1	12·0	12·6
Median number of years	10·0	10·0	10·0
Range (years)	1-40	1-40	1-40
No. of people with <1 year at site10 (28·6)	0	0	0
Duration of current service use			
Mean number of years	5·2	6·5	5·8
Median number of years	3·0	4·0	3·0
Range (years)	0-23	1-32	0-32
No. of people with <1 year at site	1	0	1

Number of months for those with <1 year at site	3	-	3
Distribution in specialist care by Marital status			
Single	77 (55·0)	71 (53·4)	148 (54·2)
Married	20 (14·3)	23 (17·3)	43 (15·8)
DeFacto	10 (7·1)	2 (1·5)	12 (4·4)
Separated	15 (10·7)	14 (10·5)	29 (10·6)
Divorced	17 (12·1)	17 (12·8)	34 (12·5)
Widowed	0 (0·0)	3 (2·3)	3 (1·1)
Other	1 (0·7)	3 (2·3)	4 (1·5)
Distribution in specialist care by Child status			
Yes	76 (54·3)	62 (46·6)	138 (50·5)
No	64 (45·7)	71 (53·4)	135 (49·5)
Number of children living at home			
0	39 (51·3)	25 (40·3)	64 (46·4)
1	11 (14·5)	21 (33·9)	32 (23·2)
2	12 (15·8)	11 (17·7)	23 (16·7)
3	2 (2·6)	3 (4·8)	5 (3·6)
4	2 (2·6)	0 (0·0)	2 (1·4)
5-6	1 (1·3)	0 (0·0)	1 (0·7)
Not listed	37 (48·7)	2 (3·2)	39 (28·2)
Distribution in specialist care by Living situation (multiple responses could be selected)			
Living with parents	21 (15·0)	26 (19·5)	47 (17·2)
Living with siblings	8 (5·7)	8 (6·0)	16 (5·9)
Living with a partner	29 (20·7)	24 (18·0)	53 (19·4)
Living with children	31 (22·1)	29 (21·8)	60 (22·0)
Living with friends	6 (4·3)	7 (5·3)	13 (4·8)
Living in shared accommodation	12 (8·6)	12 (9·0)	24 (8·8)
Living in crisis accommodation	6 (4·3)	1 (0·8)	7 (2·6)
Living in support housing	12 (8·6)	16 (12·0)	28 (10·3)
Living alone	38 (27·1)	35 (12·8)	73 (26·7)
Homeless	4 (2·9)	2 (1·5)	6 (2·2)
Other	12 (8·6)	3 (2·3)	15 (5·5)
Distribution in specialist care by Education level			
None	0 (0·0)	1 (0·8)	1 (0·4)
Primary school	4 (2·9)	2 (1·5)	6 (2·2)
Secondary school (≤ yr 10)	46 (32·9)	39 (29·3)	85 (31·1)
Secondary school (yr 11)	30 (21·4)	26 (19·5)	56 (42·1)
Secondary school (yr 12)	55 (39·3)	65 (48·9)	120 (44·0)
Not listed	5 (3·6)	0 (0·0)	5 (1·8)
Distribution in specialist care by Highest qualification			
Certificate I	2 (1·4)	2 (1·5)	4 (1·5)
Certificate II	7 (5·0)	6 (4·5)	13 (4·8)
Certificate III	26 (18·6)	21 (15·8)	47 (17·2)
Certificate IV	14 (10·0)	18 (13·5)	32 (11·7)
Diploma	18 (12·9)	12 (9·0)	30 (11·0)
Advanced Diploma	1 (0·7)	2 (1·5)	3 (1·1)
Associate Degree	0 (0·0)	1 (0·8)	1 (0·4)
Bachelor Degree	8 (5·7)	13 (9·8)	21 (7·7)
Bachelor Honours Degree	1 (0·7)	3 (2·3)	4 (1·5)
Graduate Diploma	2 (1·4)	2 (1·5)	4 (1·5)
Masters (research)	1 (0·7)	0 (0·0)	1 (0·4)
Masters (coursework)	2 (1·4)	3 (2·3)	5 (1·8)
Doctoral	1 (0·7)	1 (0·8)	2 (0·7)
Other	9 (6·4)	9 (6·8)	18 (6·6)
Not listed	48 (34·3)	40 (30·1)	88 (32·2)

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other.

Appendix 6

Stream Two Blindness findings

Process and findings

At the end of each PULSAR stream-two interview, the interviewer recorded if they thought the participant had come from a site that had received the intervention in year one (step-one) or year two (step-two). Interviewers were required to make a guess if they had no thought about the site intervention status. The null hypothesis here is that the observed proportion of correct guesses is 0·5 (i.e. half-and-half). Binomial probability theory tells us the probability of a type 1 error (the incorrect rejection of a true null hypothesis), which is shown in Table 1.

Table 6 Accuracy of interviewer guesses regarding study site in stream-two

	Number of interviews	Correct guess (number)	Correct guess (%)	Probability of Type 1 error: p value
T0	74	49	66·2%	0·004
T1	45	17	37·8%	0·964
T2	74	41	55·4%	0·208

At the first timepoint T0, the proportion of interviewers correctly guessing the intervention allocation at sites (66·2%) was significant, with p-value<0·01. On review of these results during the study the team considered they may have been influenced by the non-random order of selection options, with interviewers possibly being more likely to tick the first box (which was consistently the correct option). So for the later timepoints T1 and T2, the options for interviewers to select were randomised. At both of these timepoints, the proportion of interviewers correctly guessing the intervention allocation at sites was no different to chance (T1 was 37·8%, p-value>0·9; T2 was 55·4%, p-value>0·2).

Summary and conclusions

Assessment of blindness for stream-two interviews indicated that at T0, the proportion of interviewers correctly guessing site intervention allocation (66·2%) was significant, p<0·01. This result was possibly influenced by non-random ordering of selection options and as options were randomised for T1 and T2, the proportion interviewers correct guessing of site intervention allocation was no different from chance (T1: 37·8%, p>0·9; T2: 55·4%, p>0·2). We conclude it is unlikely through the course of the project that interview bias would represent a significant bias to findings.

Appendix 7

Main model building for stream-one QPR outcome

The model building process is shown in models 1.1 to 1.4, where model 1.4 is the final main model referred to in the manuscript. All models below have the cluster variable specified as random. Model building begins with fixed factors of timepoint and intervention status in Model 1.1. Then in model 1.2, fixed effects of sex and age group are added. Then in model 1.3, added is the fixed effect variable for sector (PMHS; MHCSS). Finally, model 1.5 has same variables as model 1.7 plus 'step' group as fixed variable.

Model 1.1 Stream-one QPR mixed model with fixed factors of timepoint and intervention status, and clusters as random.

Number of obs=942.

*** p<0·01, ** p<0·05, * p<0·1

		b	se	z	p value	ll	ul
Timepoint							
	T1	-3·59	1·13	-3·17	·002***	-5·81	-1·37
	T2	-4·78	1·60	-2·99	·003***	-7·92	-1·65
Intervention status							
	yes	4·15	1·54	2·69	·007***	1·13	7·18

Model Adjusted QPR means

		Model adj. QPR mean	Std.Err.	[95%Conf.	Interval]
Timepoint					
	T0	57·14	1·25	54·70	59·58
	T1	53·55	1·34	50·92	56·18
	T2	52·35	1·39	49·63	55·07
Intervention status*					
	0	52·25	1·46	49·37	55·12
	1	56·40	1·25	53·95	58·84

*The mean difference between treatment and control groups was 4·2 (95% Confidence interval: 1·1 – 7·2).

Model 1.2. Stream-one QPR mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs= 942**.**

*** p<0·01, ** p<0·05, * p<0·1.

		b	se	z	p value	ll	ul
Sex							
	Female	-·86	1·05	-0·82	0·414	-2·92	1·20
Age group							
	2	-·88	·86	-1·02	0·308	-2·56	0·81
	3	-3·40	·93	-3·67	0·001***	-5·21	-1·58
Timepoint							
	T1	-3·20	1·07	-2·99	0·003***	-5·29	-1·11
	T2	-4·19	1·56	-2·69	0·007***	-7·25	-1·14
Intervention status							

	Yes	3·74	1·56	2·36	0·018**	·63	6·85

Model adjusted QPR means

		QPR mean	Std.Err.	[95%Conf.	Interval]
Sex					
	Male	54·82	1·29	52·29	57·35
	Female	53·96	1·25	51·51	56·41
Age group					
	1	55·60	1·12	53·40	57·80
	2	54·72	1·22	52·34	57·10
	3	52·20	1·50	49·26	55·15
Timepoint					
	0	56·82	1·18	54·52	59·13
	1	53·63	1·38	50·93	56·33
	2	52·63	1·47	49·74	55·52
Intervention status*					
	0	52·47	1·53	49·46	55·47
	1	56·20	1·25	53·75	58·66

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0.7 – 6·8).

Model 1.3. Stream-one QPR mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=942.

*** p<0·01, ** p<0·05, * p<0·1.

		b	se	z	p value	ll	ul
Sex							
	Female	-0·818	1·05	-·78	·434	-2·87	1·23
Age group							
	2	-0·93	0·85	-1·09	·274	-2·59	0·74
	3	-3·43	0·91	-3·77	0·001***	-5·21	-1·65
Timepoint							
	T1	-3·20	1·07	-2·99	·003***	-5·30	-1·10
	T2	-4·18	1·56	-2·67	·008***	-7·24	-1·11
Intervention status							
	Yes	3·72	1·64	2·27	·023**	0·51	6·92
Sector							
	2	-1·71	2·12	-0·81	·418	-5·87	2·44

Model adjusted QPR means

		QPR mean	Std.Err.	[95%Conf.	Interval]
Sex					
	1	54·85	1·35	52·20	57·49
	2	54·03	1·22	51·63	56·42
Age group					
	1	55·68	1·13	53·46	57·90
	2	54·75	1·24	52·31	57·19
	3	52·25	1·48	49·35	55·15
Timepoint					
	0	56·87	1·203	54·51	59·23
	1	53·67	1·42	50·885	56·45
	2	52·69	1·44	49·87	55·51
Intervention status					
	0	52·53	1·57	49·45	55·60
	1	56·24	1·26	53·77	58·72
Sector					
	1	55·04	1·74	51·63	58·45
	2	53·39	1·23	50·91	55·75

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0·5 – 6·9).

Model 1.4. Stream-one QPR mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.

Number of obs=942.

*** p<0·01, ** p<0·05, * p<0·1.

		b	se	z	p value	ll	ul
Sex							
	Female	-0·81	1·04	-0·79	·431	-2·84	1·21
Age group							
	2	-0·94	0·88	-1·07	·285	-2·65	0·78
	3	-3·4	0·91	-3·78	0·001***	-5·22	-1·66
Timepoint							
	T1	-3·22	1·02	-3·16	·002***	-5·22	-1·22
	T2	-4·22	1·50	-2·82	·005***	-7·15	-1·29

		b	se	z	p value	ll	ul
Intervention status							
	Yes	3·76	1·31	2·87	·004***	1·20	6·33
Sector							
	2	-1·72	2·12	-0·81	·418	-5·87	2·43
Step group							
	2	0·15	2·08	0·07	·943	-3·93	4·22

Model adjusted QPR means

		QPR mean	Std.Err.	[95%Conf.	Interval]
Sex					
	1	54·85	1·36	52·18	57·51
	2	54·03	1·23	51·63	56·44
Age Group					
	1	55·69	1·17	53·39	57·99
	2	54·75	1·25	52·31	57·20
	3	52·25	1·48	49·34	55·16
Timepoint					
	0	56·89	1·25	54·45	59·34
	1	53·67	1·43	50·86	56·48
	2	52·67	1·38	49·97	55·37
Intervention status					
	0	52·51	1·46	49·65	55·37
	1	56·27	1·23	53·87	58·67
Sector					
	1	55·05	1·75	51·62	58·47
	2	53·33	1·24	50·91	55·75
Step					
	1	54·29	1·42	51·52	57·07
	2	54·44	1·69	51·12	57·76

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0·5 – 6·8).

Appendix 8

Interaction term model for stream-one QPR outcome

The model building process is shown in models 1.5 to 1.8, where model 1.8 is the final interaction term model referred to in the manuscript. All models below have the cluster variable specified as random. Reference groups in models can be arbitrary, and were selected based on the lowest QPR means at the specified timepoint. In Appendix 12 are same models but with the first group as the reference. Model building begins with the interaction item of time and intervention status in Model 1.5. Then in model 1.6, fixed effects of sex and age group are added. Then in model 1.7, added is the fixed effect variable for sector (PMHS; MHCSS). Finally, model 1.8 has same variables as model 1.7 plus 'step' group as fixed variable.

Model 1.5. Stream-one QPR has interaction item of time and intervention status.

Timepoint	Intervention	Coefficient	Robust Std Err.	P>|z|	95% CI	
T0	No	3·59	1·13	0·002	1·37	5·81
T1	No	Reference				
T1	Yes	4·15	1·54	0·007	1·13	7·18
T2	Yes	2·96	1·39	0·030	0·24	5·68

Number of obs=942.

Model adjusted QPR means

			Model adjusted statistics	
Timepoint	**Intervention**	**QPR raw data mean**	**QPR mean**	**95% CI**
T0	No	54·7	55·1	52·7 57·5
T1	No	51·5	51·5*	48·2 54·8
	Yes	55·3	55·6*	52·9 58·4
T2	Yes	53·9	54·4	52·1 56·8

*The mean difference between treatment and control groups at year 1 was 4·2 (95% Confidence interval: 1·1 – 7·2).

Model 1.6. Stream-one QPR has interaction item of time and intervention status, and fixed variables of age-group (<30; 30-49; 50 years and over) and sex (Male/Female).

Number of obs=942

	Coefficient	Robust Std Err.	P>|z|	95% CI
Sex				
Male		Reference		

	Female	-0·86	1·05	0·41	-2·92	1·20
Age Category						
	17-29			Reference		
	30-49	-0·88	0·86	0·31	-2·56	0·81
	50-75	-3·39	0·92	<0·001	-5·21	-1·58
Timepoint	**Intervention**					
T0	No	3·20	1·07	0·003	1·10	5·29
T1	No			Reference		
T1	Yes	3·74	1·59	0·02	0·63	6·85
T2	Yes	2·74	1·35	0·04	0·09	5·34

Model adjusted QPR means

			Model adjusted statistics		
Timepoint	**Intervention**	**QPR raw data mean**	**QPR mean**	**95% CI**	
T0	No	54·7	55·0	52·6	57·3
T1	No	51·5	51·8*	48·4	55·2
	Yes	55·3	55·5*	52·7	58·3
T2	Yes	53·9	54·5	52·1	56·9

* The mean difference between treatment and control groups at year 1 was 3·7 (95% Confidence interval: 0·6 – 6·8).

Model 1.7. Stream-one QPR Model has same variables as Model 1.6 plus sector (PMHS; MHCSS) as fixed.
Number of obs=942

			Coefficient	**Robust Std Err.**	**P>|z|**	**95% CI**	
Sex							
		Male (reference)					
		Female	-0·76	1·05	0·47	-2·81	1·29
Age Category							
		17-29 (reference)					
		30-49	-0·90	0·84	0·28	-2·55	0·75
		50-75	-3·37	0·91	0·00	-5·15	-1·59
Timepoint	**Intervention**	**Sector**					
T0	No	**PMHS**	3·56	1·47	0·02	0·68	6·43
T0	No	**MHCSS**	2·58	2·65	0·33	-2·61	7·78
T1	No	**PMHS**	(reference)				
T1	No	**MHCSS**	0·03	3·23	0·99	-6·30	6·35
T1	Yes	**PMHS**	3·99	2·23	0·07	-0·38	8·37
T1	Yes	**MHCSS**	3·33	2·64	0·31	-1·85	8·51
T2	Yes	**PMHS**	4·30	1·85	0·02	0·67	7·93
T2	Yes	**MHCSS**	0·32	2·89	0·91	-5·35	5·98

Model adjusted QPR means

					Model adjusted statistics	
Sector	**Timepoint**	**Intervention**	**QPR raw data mean**	**QPR mean**	**95% CI**	
PMHS	T0	No	55·0	55·4	51·6	59·2
	T1	No	51·1	51·8	47·0	56·7
		Yes	55·4	55·8	51·4	60·2
	T2	Yes	55·1	56·1	53·0	59·2
MHCSS	T0	No	54·3	54·4	52·6	56·3
	T1	No	52·3	51·9	47·8	55·9
		Yes	54·8	55·2	53·3	57·1
	T2	Yes	52·1	52·2	49·2	55·1

Model 1.8. Stream-one QPR Model has same variables as model 1.7 plus 'step' group as fixed variable. Model also examined interactions between four variables (sector, step, time and intervention).
Number of obs=942

		Coefficient	**Robust Std Err.**	**P>|z|**	**95% CI**	
Sex						
	Male (reference)					
	Female	-0·76	1·06	0·47	-2·85	1·32
Age Category						

			17-29 (reference)					
			30-49	-0·87	0·87	0·32	-2·56	0·83
			50-75	-3·37	0·94	0·00	-5·22	-1·52
Timepoint	**Intervention**	**Step Group**	**Sector**					
T0	No	1	PMHSPMHS	5·83	3·40	0·09	-0·84	12·49
T0	No	2	PMHS	2·36	1·50	0·12	-0·58	5·30
T1	No	2	PMHS	(reference)				
T1	Yes	1	PMHS	5·02	4·21	0·23	-3·23	13·27
T2	Yes	1	PMHS	4·48	4·37	0·31	-4·09	13·04
T2	Yes	2	PMHS	4·92	1·65	0·00	1·68	8·16
T0	No	1	MHCSS	1·99	3·32	0·55	-4·52	8·50
T0	No	2	MHCSS	4·14	3·39	0·22	-2·50	10·78
T1	No	2	MHCSS	1·29	4·11	0·76	-6·77	9·34
T1	Yes	1	MHCSS	3·11	3·23	0·34	-3·22	9·43
T2	Yes	1	MHCSS	-0·29	3·16	0·93	-6·47	5·90
T2	Yes	2	MHCSS	1·94	4·20	0·65	-6·29	10·17

Model adjusted QPR means								
Sector	**Timepoint**	**Step group**	**Intervention**	**QPR raw data mean**	**QPR mean**	**95% CI**		**Pre/post intervention diff. *Significant**
PMHS	T0	1	*No*	*57·1*	*57·2*	*54·6*	*59·7*	
	T1	1	Yes	55·4	56·4	50·9	61·8	-0·8 (z-score=0·5, p=0·64)
	T2	1	Yes	54·8	55·8	49·9	61·7	
PMHS	T0	2	*No*	*53·0*	*53·7*	*47·3*	*60·1*	
	T1	2	*No*	*51*·1	*51·3*	*45·2*	*57·5*	
	T2	2	Yes	55·4	56·2	53·0	59·5	4·9 (z-score=3·0, p=0·003)*
MHCSS	T0	1	*No*	*53·3*	*53·3*	*51·0*	*55·6*	
	T1	1	Yes	54·8	54·4	52·7	56·2	1·1 (z-score=2·7, p=0·006)*
	T2	1	Yes	51·3	51·0	50·0	52·0	
MHCSS	T0	2	*No*	*55·1*	*55·5*	*53·0*	*58·0*	
	T1	2	*No*	*52*·3	*52·6*	*47·4*	*57·8*	
	T2	2	Yes	52·7	53·2	47·8	58·7	0·7 (z-score=1·22, p=0·22)

Appendix 9

Perceived need for care findings

Instrumentation

The Perceived Need for Care Questionnaire is an interviewer administered questionnaire that in the form here used classifies seven forms of need:

1. Information about mental illness, its treatments and available services. (Information)
2. Medicine or tablets. (Medicines)
3. Counselling or talking therapy. (Counselling)
4. Practical issues such as housing or money issues. (Practical)
5. Help to improve the ability to work or use time in other ways. (Time use)
6. Help to improve the ability to look after themselves in their home. (Self-care)
7. Help to meet people for support and company (Company)

Through a branching conversationally styled question structure these needs are identified as judged by the participant to fall into four perceived need categories: no need, unmet need, partially met need, or met need.

Hypotheses

Here we examine three hypotheses, H1-H3: H 1: People in intervention as an outcome of more comprehensive assessment would identify more needs: H 2: People in intervention would be more likely to identify needs where present as met and less likely to identify them as unmet. H 3: H 2 would apply especially in more personal recovery than clinical goals areas, so here items 4-7.

Results

Table 9 Need Categories assessed with the Perceived Need for Care Questionnaire as associated with intrervention status

Perceived Need category		**PULSAR-REFOCUS Intervention status**	**No need (a)**	**Unmet need (b)**	**Partially met need (c)**	**Met need (d)**	**Proportion of all needs met (d/(b+c+d)**	**Proportion of all needs unmet b/(b+c+d)**
1	Information	Control	22	14	35	66	57·4%	12%

		Intervention	10	23	26	66	57·4%	20%
2	Medicines	Control	6	0	26	105	80·2%	0%
		Intervention	2	4	17	100	82·6%	3%
3	Counselling	Control	11	13	38	75	59·5%	10%
		Intervention	11	12	41	61	53·5%	11%
4	Practical	Control	50	35	18	34	39·1%	40%
		Intervention	42	32	15	41	46·6%	36%
5	Time use	Control	48	38	14	35	40·2%	44%
		Intervention	41	31	13	42	48·8%	36%
6	Self-care	Control	56	32	13	38	45·8%	39%
		Intervention	42	29	13	42	50·0%	35%
7	Company	Control	38	34	16	46	47·9%	35%
		Intervention	37	26	23	39	44·3%	30%

Here, given the categorical nature of the data, smaller sample sizes than for primary outcome variables, and without expectation of this part of the study being fully powered, we have kept statistical analyses very simple.
H 1 - People in intervention as an outcome of better assessment would identify more needs: Here we find people in the in intervention group identified a perceived need in 696 of 881 invitations to do so 79% while among control participants this proportion was 725/956 or 76%. A two sample test of proportions result gives a z-statistic = -1·54, p=0·0622, so in the marginal significance range of 0·05-0·10.
H 2 - People in intervention would be more likely to identify needs where present as met and less likely to identify these as unmet. Here, comparisons favour the intervention 8:5 with one tie. In 13 items, 8 favouring the intervention will occur by chance with a probability of 0·157 i.e. p=0·157 so here the probability of type I error in relation to the proposition that more needs will be identified in intervention group participants is 0·157 (here p>0·10 NS).
H 3: H 2 would apply especially in more personal recovery than clinical goals areas, so here items 4-7. Here comparisons favour the intervention 7:1. In 8 items, 7 favouring the intervention will occur by chance with a probability of 0·031. So here the p-value is 0·031 i.e. probability of type I error is 0·031 (here p< 0·05).

Conclusion

While noting the limitations of the analyses, two of the three hypotheses receive some support, one with p<0·05 and another with 0·05 < p < 0·1 while the third is favoured in terms of direction of findings, though not significantly so. Considered in the context of the overall set of measures we would rate the PNCQ findings as overall favourable for the intervention condition over controls.

Appendix 10

Stream-two models of the QPR

Model 2.1 Stream-two QPR mixed model with fixed factors of time and intervention status, and clusters as random.

Number of obs=269.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-0·50	2·79	-0·18	0·86	-5·95	4·96
T2	-2·74	3·94	-0·70	0·49	-10·45	4·98
Intervention status						
1	2·446	2·90	0·84	0·40	-3·23	8·12

Model Adjusted QPR means

	QPR mean	Std.Err.	[95%Conf.	Interval]
Timepoint				
T0	54·50	1·44	51·68	57·31
T1	54·00	1·47	51·12	56·88
T2	51·76	2·77	46·33	57·18
Intervention status				
0	52·34	2·29	47·84	56·83
1	54·78	0·93	52·96	56·60

Model 2.2 Stream-two QPR mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs=265.

	b	se	z	p value	[95%Conf.	Interval]

Sex						
Female	-2·66	1·82	-1·47	0·143	-6·22	0·90
Age group						
30-49 years	0·98	1·79	0·55	0·585	-2·53	4·49
50 years and over	-3·01	2·11	-1·43	0·153	-7·14	1·12
Timepoint						
T1	-0·88	2·75	-0·32	0·749	-6·26	4·51
T2	-2·64	3·93	-0·67	0·502	-10·34	5·06
Intervention status						
1	2·52	2·74	0·92	0·356	-2·84	7·88

Model Adjusted QPR means

	Margin	Std.Err.	[95%Conf.	Interval]
Sex				
Male	55·00	0·83	53·34	56·61
Female	52·32	1·65	49·08	55·56
Age group				
17-30 years	53·84	1·64	50·62	57·07
30-49 years	54·82	1·18	52·51	57·13
50 years and over	50·83	1·61	47·68	53·99
Timepoint				
T0	54·60	1·39	51·88	57·34
T1	53·72	1·51	50·76	56·68
T2	51·96	2·79	46·49	57·43
Intervention status or sector				
0	52·25	2·18	47·97	56·54
1	54·78	0·96	52·90	56·65

Model 2.3 Stream-two QPR mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.
Number of obs=265.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·68	1·94	-1·38	·168	-6·49	1·13
Age group						
30-49 years	·99	1·78	·55	·58	-2·50	4·48
50 years and over	-2·99	2·13	-1·41	·16	-7·16	1·18
Timepoint						
T1	-·88	2·80	-·32	·75	-6·37	4·61
T2	-2·67	4·15	-·64	·52	-10·81	5·47
Intervention status						
1	2·54	2·88	·88	·37	-3·09	8·18
Sector						
2	·23	2·07	·11	·91	-3·83	4·28

Model Adjusted QPR means

	QPR mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	55·00	·81	53·40	56·57
Female	52·31	1·74	49·00	55·71
Age group				
17-30 years	53·83	1·67	50·55	57·11
30-49 years	54·82	1·20	52·47	57·17
50 years and over	50·84	1·59	47·73	53·95
Timepoint				
T0	54·61	1·45	51·77	57·45
T1	53·73	1·49	50·81	56·64
T2	52·00	3·00	46·19	57·60
Intervention status				
0	52·24	2·27	48·00	56·69
1	54·79	·97	53·00	56·70
Sector				
1	53·38	1·82	49·81	57·00
2	53·61	·67	53·00	54·92

Model 2.4 Stream-two QPR mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.
Number of obs=265.

Step variable omitted due to collinearity == model the same as 2.3

Appendix 11

Stream-two models of the secondary outcomes

Model 3.1 Stream-two Warrick mixed model with fixed factors of time and intervention status, and clusters as random.

Number of obs=272.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-·94	2·26	-·41	·68	-5·37	3·50
T2	-1·84	3·85	-·48	·63	-9·39	5·71
Intervention status						
1	1·99	2·42	·82	·41	-2·76	6·74

Model Adjusted means

	Warrick mean	Std.Err.	[95%Conf.	Interval]
Timepoint				
T0	42·84	1·40	40·08	45·59
T1	41·89729	1·22	39·50	44·30
T2	41·00	2·77	35·57	46·43
Intervention status				
0	40·92	1·97	37·07	44·78
1	42·91	·99	40·97	44·85

Model 3.2 Stream-two Warrick mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs=268.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·08	1·271	-1·64	·10	-4·57	·41
Age group						
30-49 years	·80	1·14	·71	·48	-1·43	3·03
50 years and over	-·32	1·13	-·28	·78	-2·53	1·89
Timepoint						
T1	-1·50	2·41	-·62	·53	-6·23	3·23
T2	-2·38	4·04	-·59	·56	-10·31	5·55
Intervention status						
1	2·36	2·48	·95	·34	-2·50	7·22

Model Adjusted means

	Warrick mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	43·05	1·05	41·00	45·10
Female	40·97	1·28	38·47	43·47
Age Group				
17-30 years	41·58	1·43	38·77	44·39
30-49 years	42·38	1·21	40·01	44·76
50 years and over	41·26	1·02	39·26	43·26
Timepoint				
T0	43·23	1·51	40·27	46·18
T1	41·73	1·29	39·21	44·25
T2	40·85	2·84	35·282	46·42
Intervention status				
0	40·73	2·00	36·82	44·64
1	43·092	1·01	41·11	45·07

Model 3.3 Stream-two Warrick mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=268.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·1	1·33	-1·58	·11	-4·71	·51
Age group						
30-49 years	·815	1·13	·72	·47	-1·39	3·02
50 years and over	-·30	1·17	-·25	·80	-2·60	2·00
Timepoint						
T1	-1·51	2·46	-·61	·54	-6·33	3·31
T2	-2·42	4·15	-·58	·56	-10·55	5·72
Intervention status						
1	2·39	2·57	·93	·35	-2·66	7·43
Sector						
2	·33	2·03	·16	·87	-3·65	4·31

Model Adjusted means

	Warrick mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	43·06	1·03	41·04	45·07
Female	40·96	1·32	38·37	43·54
Age group				
17-30 years	41·57	1·46	38·72	44·41
30-49 years	42·38	1·22	39·99	44·78
50 years and over	41·27	1·02	39·28	43·26
Timepoint				
T0	43·24	1·55	40·21	46·28
T1	41·73	1·28	39·23	44·24
T2	40·82	2·90	35·14	46·51
Intervention status				
0	40·72	2·05	36·71	44·73
1	43·10	1·03	41·09	45·12
Sector				
1	41·74	1·61	38·59	44·89
2	42·07	1·12	39·87	44·27

Model 3.4 Stream-two Warrick mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.
Number of obs=268.

Step omitted due to collinearity == model the same as 3.3.

Model 4.1 Stream-two INSPIRE_S_score mixed model with fixed factors of time and intervention status, and clusters as random.
Number of obs=251.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-3·09	4·93	-·63	·53	-12·76	6·57
T2	-2·65	6·22	-·43	·67	-14·84	9·53
Intervention status						
1	1·23	5·45	·23	·82	-9·44	11·91

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
Timepoint				
T0	64·47	3·64	57·33	71·61
T1	61·38	2·74	56·00	66·76
T2	61·82	3·10	55·74	67·90
Intervention status				
0	61·74	3·32	55·24	68·25
1	62·98	2·86	57·36	68·59

Model 4.2 Stream-two INSPIRE_S_score mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.
Number of obs=248.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·61	3·19	-·819	·41	-8·86	3·64
Age group						
30-49 years	2·19	5·41	·40	·69	-8·41	12·78
50 years and over	1·30	4·89	·26	·79	-8·29	10·89
Timepoint						
T1	-3·79	4·67	-·81	·42	-12·94	5·36
T2	-3·01	6·06	-·50	·62	-14·89	8·86
Intervention status						
1	1·29	5·36	·24	·81	-9·21	11·80

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	63·88	2·64	58·72	69·05
Female	61·27	1·81	57·72	64·83
Age group				
17-30 years	60·97	4·26	52·62	69·31
30-49 years	63·15	2·30	58·65	67·66
50 years and over	62·26	2·10	58·15	66·38
Timepoint				

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
T0	64·90	3·53	57·99	71·83
T1	61·12	2·72	55·78	66·45
T2	61·89	3·07	55·88	67·90
Intervention status				
0	61·75	3·28	55·32	68·17
1	63·04	2·87	57·41	68·67

Model 4.3 Stream-two INSPIRE_S_score mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=248.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-3·49	3·03	-1·15	·25	-9·43	2·46
Age group						
30-49 years	2·43	5·19	·47	·64	-7·74	12·61
50 years and over	1·63	4·82	·34	·73	-7·82	11·09
Timepoint						
T1	-4·34	3·44	-1·26	·207	-11·08	2·40
T2	-4·21	4·50	-·93	·35	-13·04	4·62
Intervention status						
1	2·03	4·46	·45	·65	-6·72	10·78
Sector						
2	7·55	2·50	3·02	·00	2·66	12·45

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	64·37	2·19	60·09	68·66
Female	60·89	1·68	57·60	64·17
Age group				
17-30 years	60·74	4·02	52·86	68·63
30-49 years	63·18	2·09	59·24	67·11
50 years and over	62·38	2·02	58·43	66·33
Timepoint				
T0	65·46	2·32	60·91	70·01
T1	61·12	2·34	56·53	65·71
T2	61·26	2·68	56·00	66·50
Intervention status				
0	61·37	2·54	56·40	66·35
1	63·40	2·51	58·47	68·33
sector				
1	59·01	1·21	56·65	61·38
2	66·57	2·19	62·27	70·86

Model 4.4 Stream-two INSPIRE_R_score mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.

Number of obs=268.

Step omitted due to collinearity == model the same as 4.3.

Model 5.1 Stream-two INSPIRE_R_score mixed model with fixed factors of time and intervention status, and clusters as random.

Number of obs=263.

		b	se	z	p value	[95%Conf.	Interval]
Timepoint							
	T1	-1·72	4·09	-·42	·67	-9·74	6·30
	T2	1·41	5·84	·24	·81	-10·04	12·86
Intervention status							
	1	2·82	4·87	·58	·56	-6·72	12·36

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
Timepoint				
T0	74·14	3·43	21·62	80·87
T1	72·43	2·44	29·70	77·21
T2	75·55	3·06	24·73	81·54
Intervention status				

0	72·31	2·75	26·27	77·71
1	75·14	2·89	25·99	80·80

Model 5.2 Stream-two INSPIRE_R_score mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs=259.

	b	se	z	pvalue	[95%Conf.	Interval]
Sex						
Female	·78	3·10	·25	·80	-5·30	6·86
Age group						
30-49 years	2·36	5·32	·44	·66	-8·06	12·78
50 years and over	4·26	5·45	·78	·43	-6·42	14·95
Timepoint						
T1	-2·02	4·19	-·48	·63	-10·22	6·19
T2	1·10	5·74	·19	·85	-10·16	12·36
Intervention status						
1	2·45	4·86	·50	·61	-7·07	11·97

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
Sex				
Male	73·27	2·30	68·77	77·76
Female	74·05	1·93	70·27	77·83
Age Group				
17-30 years	71·35	4·41	62·70	80·00
30-49 years	73·71	2·30	69·20	78·21
50 years and over	75·61	2·03	71·63	79·59
Timepoint				
T0	74·35	3·33	67·82	80·88
T1	72·33	2·44	67·54	77·12
T2	75·45	3·12	69·33	81·57
Intervention status				
0	72·49	2·84	66·91	78·06
1	74·94	2·79	69·47	80·40

Model 5.3 Stream-two INSPIRE_R_score mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=259.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-·20	2·86	-·072	·94	-5·81	5·40
Age group						
30-49 years	2·47	4·89	·50	·61	-7·11	12·06
50 years and over	4·74	5·29	·90	·37	-5·63	15·1
Timepoint						
T1	-2·51	2·99	-·84	·40	-8·38	3·35
T2	-·44	3·43	-·13	·90	-7·16	6·27
Intervention status						
1	3·28	3·41	·96	·33	-3·39	9·97
Sector						
2	8·22	1·71	4·8	0	4·87	11·58

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
Sex				
Male	73·76	1·73	42·59	77·16
Female	73·56	1·62	45·40	76·73
Age group				
17-30 years	71·12	4·03	17·66	79·02
30-49 years	73·60	1·76	41·78	77·05
50 years and over	75·86	2·06	36·87	79·90
Timepoint				
T0	74·95	1·87	40·04	78·62
T1	72·44	1·75	41·42	75·87
T2	74·51	2·46	30·25	79·34
Intervention status				
0	72·04	1·97	36·65	75·89
1	75·33	1·86	40·45	78·98
sector				
1	70·06	·98	71·53	71·99

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
2	78·29	1·47	53·12	81·18

Model 5.4 Stream-two INSPIRE_R_score mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.
Number of obs=259

Step omitted due to collinearity == model the same as 5.3

Appendix 12

Interaction term model for stream-one QPR outcome but with reference category as the first group
Below are identical models as shown in Appendix 8. However, the below models have the first group displaying as the reference group.
The context was one of typically declining QPR scores in the groups not receiving the intervention, possibly for reasons of organisational context discussed in the manuscript. While formally speaking, chose of reference categories does not affect the outcomes when presented as model adjusted means, varying the reference categories is a way to enable inspection of differences between categories. In the context of the step-wedge design and with declining QPR scores in non-intervention groups, that T1 non-intervention scores can be seen as most indicative of a 'baseline' and this guided chose of reference category for inclusion in the manuscript. For completeness we include another set of model presentations base on T0 reference categories here.

Model 1.5b. Stream-one QPR has interaction item of time and intervention status, and cluster as random.

Timepoint	Intervention	Coefficient	Robust Std Err.	P>|z|	95% CI	
T0	No	Reference				
T1	No	-3·59	1·13	0·002	-5·81	-1·37
T1	Yes	0·56	1·24	0·651	-1·88	3·00
T2	Yes	-0·63	1·24	0·618	-3·10	1·84

Number of obs=942

Model 1.6b. Stream-one QPR has interaction item of time and intervention status, and fixed variables of age-group (<30; 30-49; 50 years and over) and sex (Male/Female) and cluster as random.
Number of obs=942.

		Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex						
	Male			Reference		
	Female	-0·86	1·05	0·41	-2·92	1·20
Age Category						
	17-29			Reference		
	30-49	-0·88	0·86	0·31	-2·56	0·81
	50-75	-3·39	0·92	<0·001	-5·21	-1·58
Timepoint	**Intervention**					
T0	No			Reference		
T1	No	-3·20	1·07	0·003	-5·29	-1·10
T1	Yes	0·54	1·26	0·669	-1·93	3·02
T2	Yes	-0·46	1·33	0·688	-2·68	1·77

Model 1.7b. Stream-one QPR. Model has same variables as Model 1.6b plus sector (PMHS; MHCSS) as fixed.
Number of obs=942.

			Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex							
		Male (reference)					
		Female	-0·76	1·05	0·47	-2·81	1·29
Age Category							
		17-29 (reference)					
		30-49	-0·90	0·84	0·28	-2·55	0·75
		50-75	-3·37	0·91	0·00	-5·15	-1·59
Timepoint	**Intervention**	**Sector**					
T0	No	**PMHS**	reference				
T0	No	**MHCSS**	-0·97	2·15	0·65	-5·19	3·24
T1	No	**PMHS**	-3·56	1·47	0·02	6·43	-0·68

T1	No	**MHCSS**	-3·53	2·83	0·21	-9·07	2·01
T1	Yes	**PMHS**	0·44	1·89	0·87	-3·26	4·14
T1	Yes	**MHCSS**	-0·23	2·14	0·92	-4·01	3·95
T2	Yes	**PMHS**	0·74	1·50	0·62	-2·18	3·67
T2	Yes	**MHCSS**	-3·24	2·43	0·18	-8·01	1·53

Model 1.8b. Stream-one QPR. Model has same variables as model 1.7b plus 'step' group as fixed variable. Model also examined interactions between four variables (sector, step, time and intervention).
Number of obs=942.

				Coefficient	**Robust Std Err.**	**P>|z|**	**95% CI**	
Sex								
			Male (reference)					
			Female	-0·76	1·06	0·47	-2·85	1·32
Age Category								
			17-29 (reference)					
			30-49	-0·87	0·87	0·32	-2·56	0·83
			50-75	-3·37	0·94	0·00	-5·22	-1·52
Timepoint	**Intervention**	**Step Group**	**Sector**					
T0	No	1	PMHS	reference				
T0	No	2	PMHS	-3·47	3·50	0·33	-10·38	3·46
T1	No	2	PMHS	-5·83	3·40	0·09	-12·49	0·84
T1	Yes	1	PMHS	-0·81	1·70	0·63	-4·14	2·53
T2	Yes	1	PMHS	-1·35	1·76	0·44	-4·81	2·11
T2	Yes	2	PMHS	-0·91	2·11	0·67	-5·05	3·24
T0	No	1	MHCSS	-3·84	1·75	0·03	-7·27	-0·41
T0	No	2	MHCSS	-1·68	1·82	0·36	-5·27	1·90
T1	No	2	MHCSS	-4·54	2·94	0·12	-10·31	1·23
T1	Yes	1	MHCSS	-2·72	1·59	0·09	-5·84	0·40
T2	Yes	1	MHCSS	-6·12	1·42	<0·001	-8·89	-3·34
T2	Yes	2	MHCSS	-3·89	3·09	0·21	-9·95	2·16

References

1. Victorian Healthcare Association. State budget submission 2015-2016. Melbourne: Victorian Healthcare Association 2015.

2. Victorian Healthcare Association. State budget submission 2017-2018. Melbourne: Victorian Healthcare Association 2017.

3. Silburn K. Recommissioning community mental health support services and alcohol and other drugs treatment services in Victoria: Report on findings from interviews with senior personnel from both sectors: La Trobe University, Australian Institute for Primary Care and Ageing; 2015.

4. Koop K. NDIS and state mental health reform: opportunities and risks for Victoria. *New Paradigm* 2014 **Summer**.

5. Mental Health Act 2014.

6. Klein JH, Young T. Health care: a case of hypercomplexity? *Health Systems* 2015; **4**(2): 104-10.

7. Long KM, McDermott F, Meadows GN. Being pragmatic about healthcare complexity: our experiences applying complexity theory and pragmatism to health services research. *BMC Med* 2018; **16**(1): 94.

8. Long KM, Mc Dermott F, Meadows GN. Mental health services management and simulation modelling: A project evaluation drawing on pragmatism, complexity theory and the critical incident approach. (unpublished).

Figure 1

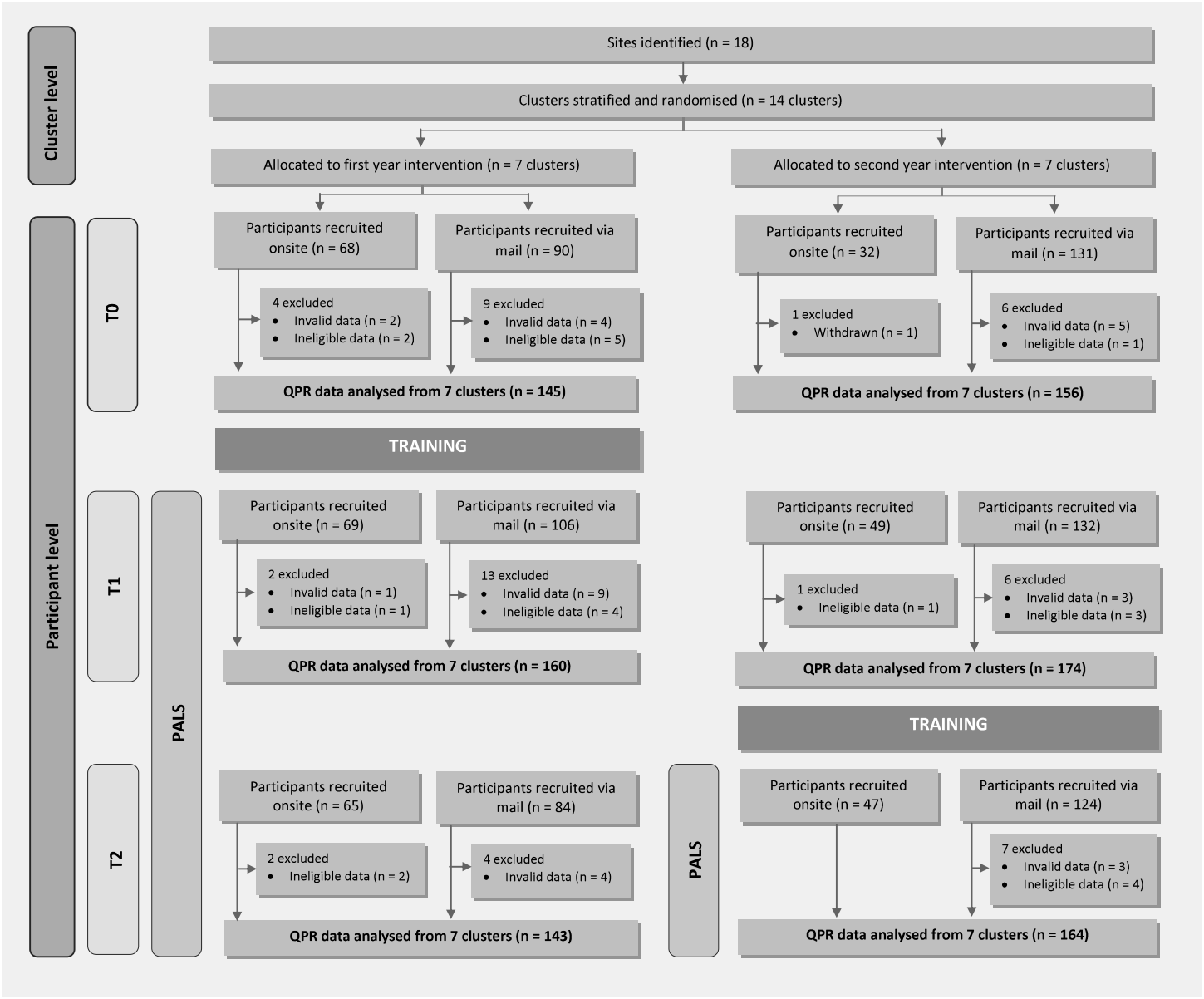

Cluster level
Sites identified (n = 18)
Clusters stratified and randomised (n = 14 clusters)
Allocated to first year intervention (n = 7 clusters)
Allocated to second year intervention (n = 7 clusters)
Participant level
T0
Participants recruited onsite (n = 68)
Participants recruited via mail (n = 90)
4 excluded
Invalid data (n = 2)
Ineligible data (n = 2)
9 excluded
Invalid data (n = 4)
Ineligible data (n = 5)
QPR data analysed from 7 clusters (n = 145)
Participants recruited onsite (n = 32)
Participants recruited via mail (n = 131)
1 excluded
Withdrawn (n = 1)
6 excluded
Invalid data (n = 5)
Ineligible data (n = 1)
QPR data analysed from 7 clusters (n = 156)
TRAINING
T1
PALS
Participants recruited onsite (n = 69)
Participants recruited via mail (n = 106)
2 excluded
Invalid data (n = 1)
Ineligible data (n = 1)
13 excluded
Invalid data (n = 9)
Ineligible data (n = 4)
QPR data analysed from 7 clusters (n = 160)
Participants recruited onsite (n = 49)
Participants recruited via mail (n = 132)
1 excluded
Ineligible data (n = 1)
6 excluded
Invalid data (n = 3)
Ineligible data (n = 3)
QPR data analysed from 7 clusters (n = 174)
TRAINING
T2
PALS
Participants recruited onsite (n = 65)
Participants recruited via mail (n = 84)
2 excluded
Ineligible data (n = 2)
4 excluded
Invalid data (n = 4)
QPR data analysed from 7 clusters (n = 143)
PALS
Participants recruited onsite (n = 47)
Participants recruited via mail (n = 124)
7 excluded
Invalid data (n = 3)
Ineligible data (n = 4)
QPR data analysed from 7 clusters (n = 164)

Figure 2

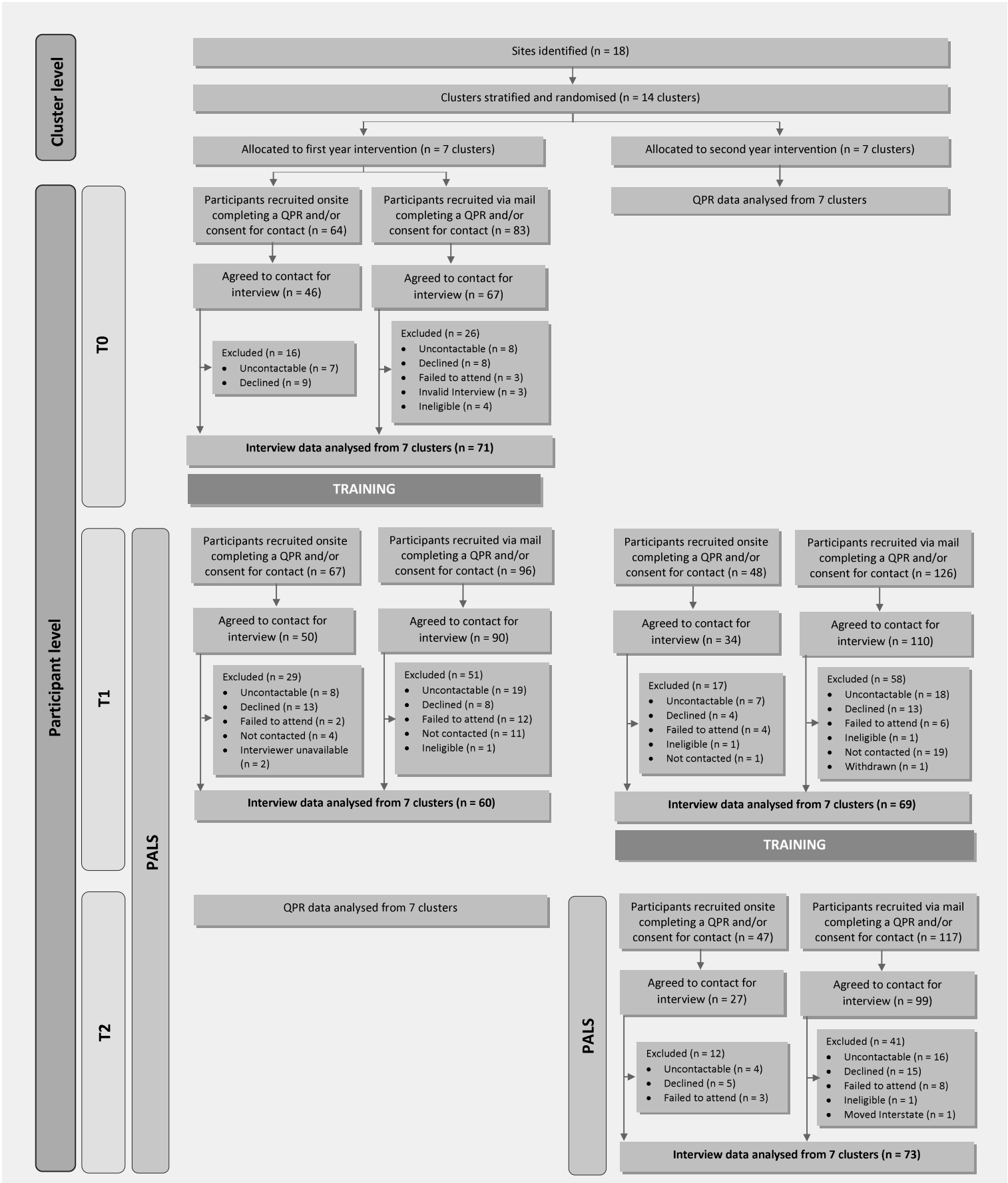

Cluster level
Participant level
T0
T1
T2
PALS
Sites identified (n = 18)
Clusters stratified and randomised (n = 14 clusters)
Allocated to first year intervention (n = 7 clusters)
Allocated to second year intervention (n = 7 clusters)
QPR data analysed from 7 clusters
Participants recruited onsite completing a QPR and/or consent for contact (n = 64)
Participants recruited via mail completing a QPR and/or consent for contact (n = 83)
Agreed to contact for interview (n = 46)
Agreed to contact for interview (n = 67)
Excluded (n = 16)
Uncontactable (n = 7)
Declined (n = 9)
Excluded (n = 26)
Uncontactable (n = 8)
Declined (n = 8)
Failed to attend (n = 3)
Invalid Interview (n = 3)
Ineligible (n = 4)
Interview data analysed from 7 clusters (n = 71)
TRAINING
Participants recruited onsite completing a QPR and/or consent for contact (n = 67)
Participants recruited via mail completing a QPR and/or consent for contact (n = 96)
Agreed to contact for interview (n = 50)
Agreed to contact for interview (n = 90)
Excluded (n = 29)
Uncontactable (n = 8)
Declined (n = 13)
Failed to attend (n = 2)
Not contacted (n = 4)
Interviewer unavailable (n = 2)
Excluded (n = 51)
Uncontactable (n = 19)
Declined (n = 8)
Failed to attend (n = 12)
Not contacted (n = 11)
Ineligible (n = 1)
Interview data analysed from 7 clusters (n = 60)
Participants recruited onsite completing a QPR and/or consent for contact (n = 48)
Participants recruited via mail completing a QPR and/or consent for contact (n = 126)
Agreed to contact for interview (n = 34)
Agreed to contact for interview (n = 110)
Excluded (n = 17)
Uncontactable (n = 7)
Declined (n = 4)
Failed to attend (n = 4)
Ineligible (n = 1)
Not contacted (n = 1)
Excluded (n = 58)
Uncontactable (n = 18)
Declined (n = 13)
Failed to attend (n = 6)
Ineligible (n = 1)
Not contacted (n = 19)
Withdrawn (n = 1)
Interview data analysed from 7 clusters (n = 69)
TRAINING
QPR data analysed from 7 clusters
Participants recruited onsite completing a QPR and/or consent for contact (n = 47)
Participants recruited via mail completing a QPR and/or consent for contact (n = 117)
Agreed to contact for interview (n = 27)
Agreed to contact for interview (n = 99)
Excluded (n = 12)
Uncontactable (n = 4)
Declined (n = 5)
Failed to attend (n = 3)
Excluded (n = 41)
Uncontactable (n = 16)
Declined (n = 15)
Failed to attend (n = 8)
Ineligible (n = 1)
Moved Interstate (n = 1)
Interview data analysed from 7 clusters (n = 73)

Figure 3

Figure 3a: Model adjusted mean QPR scores over time: Public Mental Health Services (PMHS)

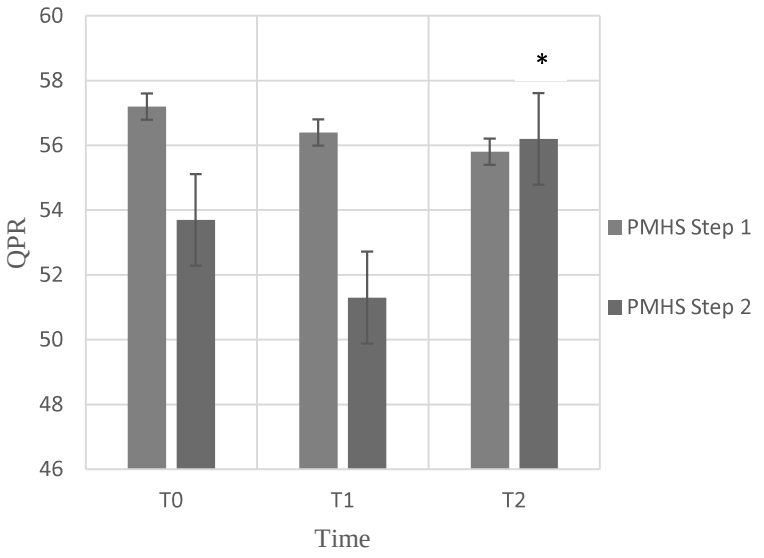


Figure 3b: Model adjusted mean QPR scores over time: Mental Health Community Support Services (MHCSS)

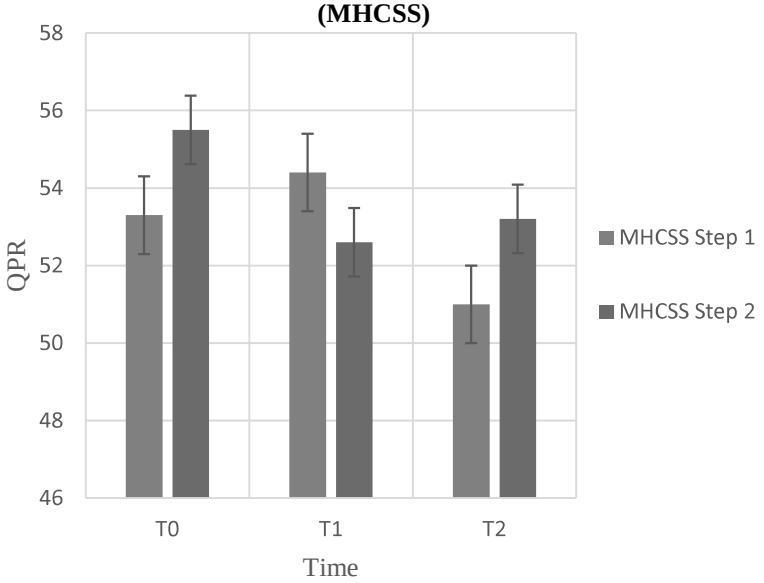
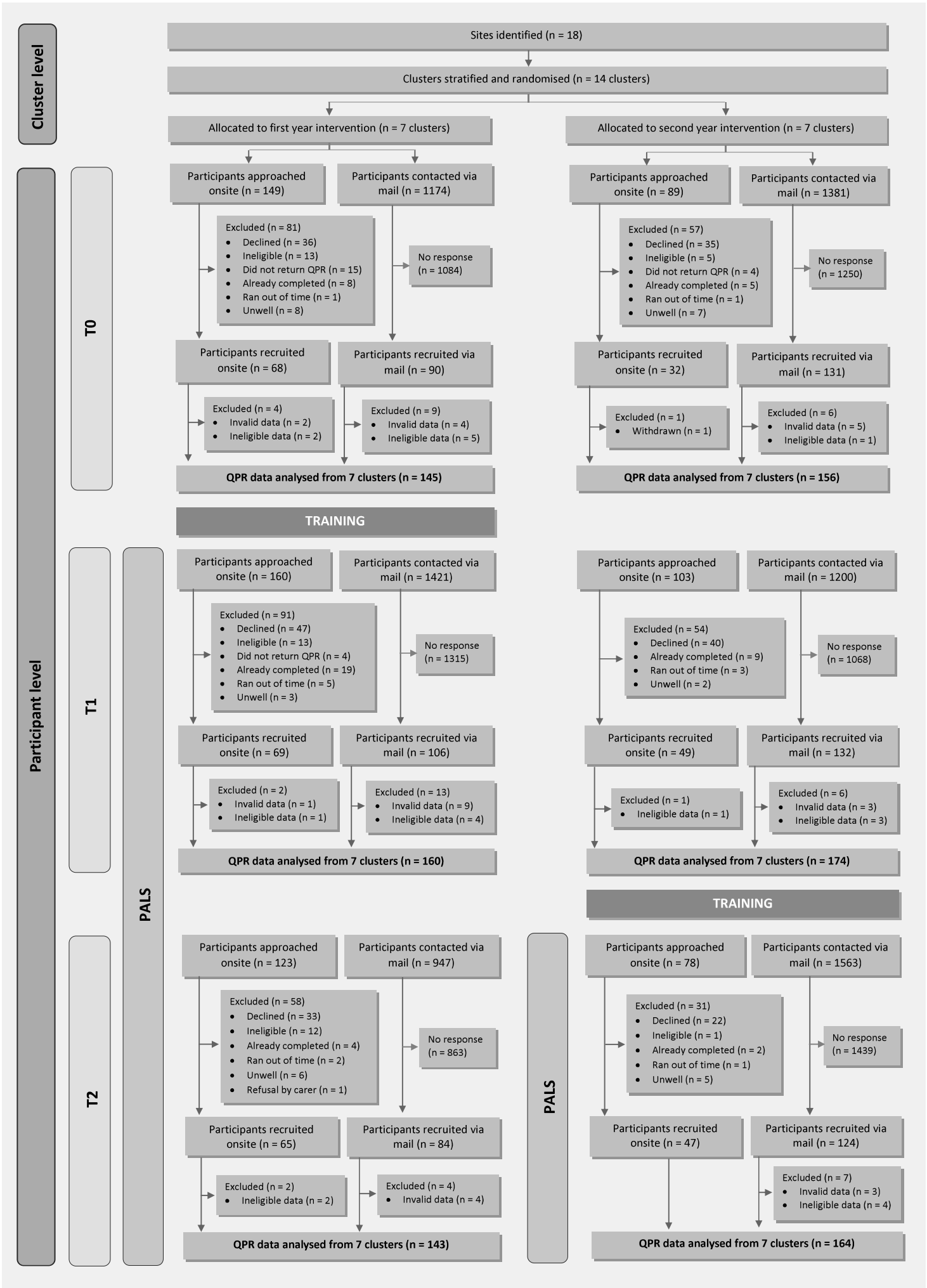


Figure S1


Cluster level
Participant level
Sites identified (n = 18)
Clusters stratified and randomised (n = 14 clusters)
Allocated to first year intervention (n = 7 clusters)
Allocated to second year intervention (n = 7 clusters)
T0
Participants approached onsite (n = 149)
Participants contacted via mail (n = 1174)
Excluded (n = 81)
Declined (n = 36)
Ineligible (n = 13)
Did not return QPR (n = 15)
Already completed (n = 8)
Ran out of time (n = 1)
Unwell (n = 8)
No response (n = 1084)
Participants recruited onsite (n = 68)
Participants recruited via mail (n = 90)
Excluded (n = 4)
Invalid data (n = 2)
Ineligible data (n = 2)
Excluded (n = 9)
Invalid data (n = 4)
Ineligible data (n = 5)
QPR data analysed from 7 clusters (n = 145)
Participants approached onsite (n = 89)
Participants contacted via mail (n = 1381)
Excluded (n = 57)
Declined (n = 35)
Ineligible (n = 5)
Did not return QPR (n = 4)
Already completed (n = 5)
Ran out of time (n = 1)
Unwell (n = 7)
No response (n = 1250)
Participants recruited onsite (n = 32)
Participants recruited via mail (n = 131)
Excluded (n = 1)
Withdrawn (n = 1)
Excluded (n = 6)
Invalid data (n = 5)
Ineligible data (n = 1)
QPR data analysed from 7 clusters (n = 156)
TRAINING
PALS
T1
Participants approached onsite (n = 160)
Participants contacted via mail (n = 1421)
Excluded (n = 91)
Declined (n = 47)
Ineligible (n = 13)
Did not return QPR (n = 4)
Already completed (n = 19)
Ran out of time (n = 5)
Unwell (n = 3)
No response (n = 1315)
Participants recruited onsite (n = 69)
Participants recruited via mail (n = 106)
Excluded (n = 2)
Invalid data (n = 1)
Ineligible data (n = 1)
Excluded (n = 13)
Invalid data (n = 9)
Ineligible data (n = 4)
QPR data analysed from 7 clusters (n = 160)
Participants approached onsite (n = 103)
Participants contacted via mail (n = 1200)
Excluded (n = 54)
Declined (n = 40)
Already completed (n = 9)
Ran out of time (n = 3)
Unwell (n = 2)
No response (n = 1068)
Participants recruited onsite (n = 49)
Participants recruited via mail (n = 132)
Excluded (n = 1)
Ineligible data (n = 1)
Excluded (n = 6)
Invalid data (n = 3)
Ineligible data (n = 3)
QPR data analysed from 7 clusters (n = 174)
TRAINING
T2
Participants approached onsite (n = 123)
Participants contacted via mail (n = 947)
Excluded (n = 58)
Declined (n = 33)
Ineligible (n = 12)
Already completed (n = 4)
Ran out of time (n = 2)
Unwell (n = 6)
Refusal by carer (n = 1)
No response (n = 863)
Participants recruited onsite (n = 65)
Participants recruited via mail (n = 84)
Excluded (n = 2)
Ineligible data (n = 2)
Excluded (n = 4)
Invalid data (n = 4)
QPR data analysed from 7 clusters (n = 143)
PALS
Participants approached onsite (n = 78)
Participants contacted via mail (n = 1563)
Excluded (n = 31)
Declined (n = 22)
Ineligible (n = 1)
Already completed (n = 2)
Ran out of time (n = 1)
Unwell (n = 5)
No response (n = 1439)
Participants recruited onsite (n = 47)
Participants recruited via mail (n = 124)
Excluded (n = 7)
Invalid data (n = 3)
Ineligible data (n = 4)
QPR data analysed from 7 clusters (n = 164)

Positive Changes Associated With a Recovery-Oriented Mental Health Care Training Intervention in the REFOCUS-PULSAR Specialist Care Cluster Stepped-Wedge Randomised Controlled Trial

Graham Meadows*, Lisa Brophy, Frances Shawyer, Joanne C. Enticott, Ellie Fossey, Christine D. Thornton, Penelope Weller, Elisabeth Wilson-Evered, Vrinda Edan, Mike Slade

Author details

- Graham Meadows MD, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia; Melbourne School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia; Monash Health, Melbourne, VIC, Australia.
- Lisa Brophy PhD, Centre for Mental Health, Melbourne School of Population and Global Health, University of Melbourne, Parkville, VIC, Australia. Mind Australia, Heidelberg, VIC, Australia. School of Allied Health, La Trobe University.
- Frances Shawyer PhD, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia.
- Joanne C. Enticott PhD, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health; Department of General Practice, School of Primary and Allied Health Care, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia; Centre of Research Excellence in Suicide Prevention, Brain and Mind Centre, University of Sydney, Camperdown, NSW, Australia.
- Ellie Fossey PhD, Department of Occupational Therapy, School of Primary and Allied Health Care, Monash University - Peninsula Campus, Frankston, VIC, Australia.
- Christine D. Thornton BA, Ermha Ltd, Dandenong, VIC; Core Advanced Technologies Limited, Worcestershire, UK; Melbourne School of Law, University of Melbourne, Parkville, VIC, Australia.
- Penelope June Weller PhD, Graduate School of Business and Law, RMIT University, Melbourne, VIC, 3001 Australia.
- Elisabeth Wilson-Evered PhD, Institute of Health and Sport, Victoria University, Melbourne, VIC, Australia.
- Vrinda Edan, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC, Australia.
- Mike Slade PhD, School of Health Sciences, Institute of Mental Health, University of Nottingham, Triumph Road, Nottingham, NG7 2TU, UK.

*Corresponding author:
Professor Graham Meadows, Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Dandenong Hospital, 126 - 128 Cleeland St, Dandenong Victoria 3175, Australia. Tel: (61 3) 9902 9696; Fax: (61 3) 9902 9900;
graham.meadows@monash.edu

Formatted: Right: -1.15 cm

This project was undertaken through Monash University.

Summary

Background: Recovery-oriented practice promotes individual~~recognises people's~~ strengths and recovery potential~~. even in the context of ongoing symptoms.~~ PULSAR, adapting ~~adapted~~ the UK-developed REFOCUS recovery-oriented staff intervention for Australian use,~~use in Australia across sectors. This study~~ aimed to establish whether consumers accessing ~~specialist~~ mental health services where staff had~~services who~~ received the REFOCUS-PULSAR intervention showed~~reported~~ increased ~~personal~~ recovery compared to consumers ~~who did not.~~

~~Methods:~~ of non-~~The~~ intervention services~~was implemented within a~~.

Methods: A pragmatic two-step ~~complete~~ stepped-wedge randomised controlled trial at 18 sites grouped into 14 clusters across Public Mental Health Services (PMHS) and Mental Health Community Support Services (MHCSS). Staff training was refined between step-one and step-two. The primary (stream-one)~~The primary~~ outcome measure was the Questionnaire about the Process of Recovery (QPR) with cross-sectional data collected across three time-points. Stream-two, with two data-collection points, included five outcome-measures and five experience-measures. This trial is registered with ANZCTR, number ACTRN12614000957695.

Findings: Half of the available staff were trained (190), with substantial staff turnover across the three organisations (27-47%). Between 2014 and 2017, 942 stream-one consumer participants were recruited over three time-points (T0: 301; T1: 334; T2: 307) with 273 stream-two participants recruited at intervention-related time-points. (baseline: 140, follow-up: 133). The main mixed~~.~~ ~~Mixed~~-effects model~~modelling~~ showed a small significant overall positive intervention stream-one effect of 3·7 (95% Confidence interval: ~~($p<$~~0.5 – 6·8). Examining interactions, the~~·05);~~ mean difference between intervention and control groups at year-one also ~~1~~ was 3·7 (95% Confidence interval: 0·6 – 6·8); findings were strongest for PMHS step-two. Stream-two findings of small effects, typically below study power threshold, favoured~~4·2 (95% Confidence interval: 1·1 – 7·2). Pooled data show the year 1 non-intervention clusters mean decreasing 3·4 points compared to baseline, then improving back to baseline in year 2 once all clusters had received~~ the intervention condition for all but one measure. ~~. Clusters receiving the intervention in year 1 did not exhibit this decrease. PMHS findings were similar to the overall result, whereas the MHCSS mean remained low in year two in the context of major organisational and commissioning changes.~~

Formatted: Right: -1.15 cm

Interpretation: The ~~fully developed~~ REFOCUS-PULSAR intervention showed modest but distinct effectiveness~~was effective~~ in promoting recovery-oriented practice across sectors~~, over 1 year, and 2 year sustainability was suggested in PMHS.~~

Funding: Victorian Government Mental Illness Research Fund.

Key words: Recovery, Recovery-oriented practice, Specialist mental health services, Mental Health, Training, Psychiatry, Cluster Randomised Controlled Trial, Pragmatic trial, Health services research, Complex intervention, Questionnaire about the Process of Recovery (QPR).

Research in context

Evidence before the study

Searching PsycINFO, Medline and CINAHL, ~~without language restrictions,~~ for articles published in English between 1 January 2007 and 31 July 2017, given the development and evaluation of approaches to implementing recovery-oriented practices is relatively recent. The search ~~January 2018, with a~~ strategy included the following search terms: [Mental Health/ OR "mental health" OR Mental Health Services/ OR "mental health service*"] AND [Recovery/ OR recover*],~~combining Boolean key term operators of: recovery oriented practice; community mental health services; implementation; and staff training,~~ then identification of further relevant articles from reference lists of key papers, author searches and citation searches in Google Scholar. We selected articles if they were set in community mental health services and included data related to staff views, staff-related outcomes or consumer-related outcomes in the context of staff training in recovery-oriented practice (ROP) and/or implementation of ROP to promote and support personal recovery. This~~,~~ identified 16~~21~~ relevant studies typically~~. Typically~~ assessing staff-related outcomes after recovery-oriented training programs. While only REFOCUS had been evaluated using a randomised controlled trial design, these studies generally suggest that recovery-oriented training~~various programs~~ improved staff knowledge~~, self-efficacy~~ and attitudes towards recovery and improved self-efficacy towards providing recovery-oriented care, with a recurrent theme that the ~~institutional and~~ organisational culture of the service ~~implementation~~ setting, and the provision of follow-up coaching appear to be~~ is an~~ important determinants~~determinant~~ of implementation~~intervention~~ success. Apart from the REFOCUS trial published in 2015, no others have reported whether consumer outcomes were improved by~~from~~ these interventions.

Formatted: Right: -1.15 cm

Added value of the study

The REFOCUS-PULSAR staff training intervention, adapted for Australian service settings from the REFOCUS package and based on the CHIME (Connectedness, Hope, Identity, Meaning, and Empowerment) conceptual framework of personal recovery, was examined through a stepped-wedge randomised controlled trial with quantitative assessment of effect on consumer-rated experience of recovery. Positive findings for intervention effect in the study provide evidence that the REFOCUS-PULSAR intervention as developed and implemented in this study brought about modest improvements in consumer-rated recovery for people using the involved services. The findings also suggest possible improvements in clinical recovery and experience of service.

Implications of all the available evidence

Training health-care workers to deliver recovery-oriented care using the REFOCUS materials developed over time and adapted to local settings can positively influence the process of personal recovery for consumers.

Formatted: Right: -1.15 cm

Introduction

Developing evidence around recovery orientation

The construct of recovery now commonly used in mental health care has roots in consumer perspectives[1] and may be distinguished from other conceptualisations by reference to *personal* rather than *clinical* recovery.[2] Recovery-oriented practice (ROP) involves clinical and other staff facilitating a change process through which individuals who have been diagnosed with mental illness are supported to live a self-directed life and to strive to reach their full potential.[3] Promoting recovery within mental health services is well established in mental health policy internationally[4] and in Australia[5] where this study is set. However, the practice lags behind policy: service-level intervention is required to effectively implement practices through which mental health professionals employ skills, values, attitudes and behaviours that support individuals in their personal recovery.[6] The past decade has seen the development of a number of recovery-oriented training programs, such as REFOCUS[6] and THRIVE[7] in the UK, the Collaborative Recovery Model[8,9] in Australia and Person-Centred Recovery Planning[10] in the USA. They typically emphasize the use of coaching and person-centred, strengths focused and collaborative processes for supporting service users in their recovery. A useful reference framework for the work on training interventions may be Kirkpatrick's four levels of learning evaluation: K1-reaction, K2-learning, K3-behaviour and K4-results.[11] The literature is strongest on levels 1 and 2, with few programs having evidence at either level 3 or level 4. Typically work at level 4 has not had the strength of evidential value that goes with RCT methods so there is a need for further evidence at this level. Evidence of the effectiveness of these interventions to promote ROP is required across settings, so that they might be adopted with some confidence by services working towards these policy goals.

~~The construct of recovery now commonly used in mental health care has roots in consumer perspectives[1] and may be distinguished from other conceptualisations by reference to *personal* rather than *clinical* recovery.[2] Recovery-oriented practice (ROP) involves clinical and other staff facilitating a change process through which individuals who have been diagnosed with mental illness are supported to live a self-directed life and to strive to reach their full potential.[3] Promoting recovery within mental health services is well established in mental health policy internationally[4] and in Australia[5] where this study is set. Yet, there is a need for interventions to promote ROP that have evidence of effectiveness across settings and so might be adopted with some confidence by services working towards these policy goals.~~

Formatted: Right: -1.15 cm

From REFOCUS to PULSAR – a developmental trajectory

REFOCUS is a staff training intervention developed and trialled in the UK.[2,6,12] In a developmental process informed by the theory of planned behaviour,[13] ~~that has been developed and trialled in the UK.[2,6,7] In a developmental process informed by the theory of planned behaviour,[8]~~ working towards changing both what practitioners might do with consumers of mental health services (consumers) and how they might do it,[14,9] the REFOCUS intervention came to include, as elements of a team-based training intervention for community mental health teams in England, three working practices of 'understanding values and treatment practices', 'working to strengths', and 'supporting goal striving'. So, the REFOCUS intervention was designed to promote recovery through changes in staff and team skills, knowledge, behaviour, values, and relationships with consumers.[2,12] ~~[2,6]~~

In a large-scale cluster Randomised Controlled Trial (cRCT), ~~the~~ outcomes of usual care plus REFOCUS were compared with usual care only ~~(control group)~~ in 27 community mental health teams delivering services to adult consumers with psychotic disorders. In ~~the~~ primary analyses, personal recovery assessed using the consumer-rated Questionnaire about the Process of Recovery (QPR)[15] ~~)[10]~~ did not differ between the ~~REFOCUS~~ intervention group and controls. Secondary ~~While secondary~~ analyses suggested ~~that~~ higher team participation was associated with higher staff-reported recovery-promoting behaviour and improved QPR. Possible ~~, possible~~ reasons advanced for the negative primary analyses that might be modifiable in subsequent work included the following issues:[6,7] ~~,~~

1. The REFOCUS recruitment protocol and criteria meant that, on average, consumer participants had been using mental health services for ≥ ~~over~~ 15 years, suggesting the possibility of entrenched ways of relating to services, and problems that may take longer than one year to change.
2. Participant attrition, ~~was~~ higher than anticipated in this 12-month longitudinal study (26% vs 7%), ~~%)~~ resulting in a reduction in planned statistical power.
3. Inclusion of adaptive design principles[16,17] ~~a pilot phase~~ might be advantageous.
4. Future designs might either use a homogenous team-type or stratification by team characteristics.
5. Transition to ROP might require organisation-wide rather than team-level strategies.

The 'Principles Unite Local Services Assisting Recovery' (PULSAR) work program was based in Victoria, Australia. The REFOCUS team advised on project development enabling PULSAR, four

Formatted: Right: -1.15 cm

years behind REFOCUS in development and implementation, to benefit from lessons learned during REFOCUS. Changes to the intervention included adjustments to the REFOCUS materials to enhance relevance to the local setting and to incorporate developments made in the course of the REFOCUS work after the REFOCUS manual[18] was concluded for~~manual[11] was concluded for REFOCUS~~ study use. The intervention here is referred to as "REFOCUS-PULSAR" (shortened to "PULSAR" in the protocol paper and local implementation[19]) since while it was developed for the PULSAR study,[19] it drew heavily~~implementation[12]) since while it was developed for the PULSAR study,[12] it was heavily based~~ on REFOCUS materials.

The research approach,[19] chosen based on addressing issues 1-5 above,~~The research approach,[12] chosen based on addressing issues 1-5 above, is outlined here with that context as background to the methods section. It~~ involved adoption of a specific cRCT variant involving Stepped-Wedge intervention allocation (a cRCT-SW) where all study sites receive the intervention but time of intervention is allocated randomly, here according to two 'steps', step-one and step-two. Since those people who may benefit most from ROP in relation to personal recovery may also experience clinical recovery and so be discharged earlier from ~~not stay very long with~~ treating services, sampling based on people with long-term service tenure ~~with those services~~ may bias against positive findings as noted in point 1 above. Hence, the PULSAR design primary~~, rather than following individuals longitudinally through the three years of the study period, involves~~ recruitment strategy recruited independently~~of a different sample~~ at three~~each~~ time-points~~point~~ (baseline: T0; year 1: T1; year 2: T2) with~~). It is necessary then to maintain~~ tight control on consistency of recruitment processes so that ~~any~~ sampling bias is minimised as a source of systematic error in findings related to intervention effect across time-points. The cRCT-SW research design with repeated cross-sectional recruitment, then, carried possible advantages for point 1-2 above. The two~~2~~-year two-step stratified cRCT-SW approach promised greater possibility for progressive refinement of the training intervention through experience, providing some response to point 3. Randomisation in this study was stratified by team type addressing point 4 above. The design also went a small way to address point 5 above since in the later stage of the stepped-wedge design the implementation was in effect organisation-wide across community services. ~~The PULSAR work program also included face-to-face interviews in two smaller parallel studies, one cross-sectional and the other longitudinal, each with different samples[12], findings from which will be presented in later papers. The outcome measure collected across all time-points, identified as primary outcome in the protocol paper[12] and on which the study power calculations were based, was the QPR. These QPR findings are the subject of this paper.~~

Formatted: Right: -1.15 cm

Aims and hypotheses

The aim of this pragmatic cluster stepped-wedge randomised controlled trial was to evaluate the effectiveness of the REFOCUS-PULSAR staff ROP training intervention for improving the experience of personal recovery as reported by consumers using repeated cross-sectional samples. The primary hypothesis was~~. It was hypothesised~~ that consumers in the REFOCUS-PULSAR post-intervention clusters would experience significantly greater personal recovery ~~as measured by QPR~~ compared to consumers accessing other mental health services that at relevant time-points within the cRCT-SW had not received the intervention. We also investigated change in clinical recovery and experience of the services.

Methods

Setting

Participating services were providers of~~involved with providing~~ mental health care to~~ over one million~~ people living in the catchment area of a large Public Mental Health Service (PMHS) in Victoria, Australia. The~~ catchment~~ area ranges from a relatively affluent coastal city area to the most socio-economically disadvantaged and culturally-diverse area in metropolitan Melbourne and includes a semi-rural growth-corridor. In Victoria, state-run area-based and block-funded PMHSs~~public sector specialist clinical mental health services~~, typically accessed by people with more severe mental illnesses, ~~are block-funded. Specialist mental health services~~ include ~~area-based~~ clinical services comprising a range of teams and service types. Here are included~~, including~~ inpatient units, community-based residential rehabilitation, continuing care, and community treatment teams. Acute or longer-term Residential care~~, whether acute or longer-term,~~ is typically provided in units of around 25 beds. Caseloads in community services vary from around 10 in Mobile Support and Treatment Services (MSTS) to 25-35 in many community clinics while typical length of care with a particular team may vary between a few days with~~as in~~ Crisis Assessment and Treatment Teams (CATTS) to several years with~~as in~~ MSTS and Community Care Units (CCUs). Mental health care funded by the Victorian government also includes substantial investment in the Mental Health Community Support Services (MHCSS) sector which,~~),~~ run by non-government organisations, provides~~which provide~~ residential and outreach psychosocial support. ~~Government-funded provision of care to the population in need of mental health care for serious mental illness is thus provided by a combination of PMHS and MHCCS.~~

Formatted: Right: -1.15 cm

Within this setting, the temporal context for the work through 2014-2016 included events~~what in organisational terms could be framed as 'Critical Incidents'~~ worthy of some comment – details on these are provided in Appendix 1).~~:~~

- ~~The funding environment for public health services in Victoria under the Liberal administration 2010 to 2015 received significant criticism as negative.[13] While the Labor administration that followed has been better reviewed for its support of healthcare,[14] there was limited time for the actions of this new administration to flow through into changes in work context during the project timespan.~~
- ~~In 2014 the Victorian State Government introduced a new Mental Health Act,[15] replacing the 1986 Act; but only limited training was offered in preparation for the significant changes in practice required for compliance.~~
- ~~In 2015 the Victorian State Government introduced major reforms of MHCSS which presented substantial challenges for MHCSS organisations involved in this study.[16] This also was an element of the preparation for transition to an individualised funding model under the new National Disability Insurance Scheme, a transition commonly referred to as the biggest change in health services funding in Australia since its current National Health Insurance scheme (Medicare) was introduced in 1984.[17]~~

~~Randomisation and masking~~

The State-funded organisations that operated in the catchment were the major PHMS and two organisations from the MHCSS sector. Specialist care sites or teams within these organisations were identified by the PMHS and MHCSS service partners then approached; all agreed to participate.

Design overview

Specialist-care PULSAR project data collection from consumers included three streams. Stream-one, a cross-sectional complete step-wedge cRCT with self-administered instruments, collected QPR and demographic data. The QPR, identified as the primary outcome[19] was the basis for stream-one power calculations. Stream-two, a cross-sectional pre- and post-intervention incomplete step-wedge cRCT, involved face-to-face interviews with a subset of stream-one participants. Stream-three, a longitudinal incomplete step-wedge cRCT involving consumers from Stream-two with diagnosed psychotic disorders, did not achieve adequate recruitment targets and is not reported here.

Formatted: Right: -1.15 cm

Participants

Staff

Participating teams' members were eligible to receive the PULSAR training intervention if they were working part-time or full-time in a direct service role and had an active caseload with consumers being recruited for the evaluation. Casually employed staff or those also working in a non-intervention site at the time of training were ineligible.[19]

Consumers Stream-one

Eligible consumers were: receiving care from a participating cluster with contact in the three months prior to data collection; aged 18-75; able to provide informed consent; proficient in English; and not imprisoned. Eligibility screening, conducted by administration and clinical staff at participating organisations, used detailed instructions provided by the research team. A letter sent to all eligible consumers from participating sites invited completion and return of a demographics/QPR survey form and a contact details/consent to be contacted for a face-to-face interview form. An AUD$10 shopping voucher was sent to participants for returned surveys where contact details were provided. Additional recruitment strategies to encourage consumer response to the mailouts were utilized according to site need. Strategies included, for example, having researchers, including consumer researchers, speak about PULSAR at participating sites and use of PULSAR-branded publicity materials.[19] Through an active quality assurance process monitoring recruitment, and because this was important to the design, the balance of recruitment between onsite recruitment and mailout approaches was kept as consistent as possible across timepoints and clusters. Decisions on whether or not to repeat bulk mailouts for given clusters or continue onsite recruitment were based on a weekly review of QPR numbers by recruitment method by cluster and taking into consideration the need to also recruit sufficient numbers for face-to-face interviews. Time spent recruiting at T1 and T2 at a given cluster was matched to T0 activity at the same cluster and only adjusted if necessary to match the number of QPRs collected via this method.

Consumers stream-two

Consumers were eligible for stream-two and recruited by phone, email or letter if they had provided contact details, consent to be contacted for this purpose and were at the pre- or post-phase of an active intervention site at the time of recruitment.

Randomisation and masking

Eighteen ~~mental health~~ care ~~-~~ delivery teams, grouped into 14 clusters to enable adequate recruitment in the context of some smaller teams, were classified into seven strata. Team characteristics varied so

Formatted: Right: -1.15 cm

~~In the context of substantial variations in the nature and intensity of care typically delivered through the different teams, the~~ strata groupings included teams similar in specified function.~~that were relatively homogenous regarding the specifications of their care delivery.~~ Within PMHS these were: CATTS (x3 teams; two smaller teams grouped into one cluster) and MSTS (x2 teams); CCUs (x2; grouped with MSTS, being smaller teams and introduced earlier as having shared focus on long term intensive work with people with more complex needs); Community Mental Health Services/Continuing Care Teams (x4). The remaining stratum included services delivered by two participating MHCSS, here designated MHCSS-1 and MHCSS-2. These were: Prevention and Recovery Care services (PARCs; x4) delivering short-term, subacute, residential recovery-oriented care; and Community Outreach Services (x3; two from the one organisation grouped into one cluster).

Stratified randomisation was used to allocate clusters to receive the intervention in either step-one or step-two using an online Research Randomiser with randomisation keys corresponding to the seven strata and allocation of clusters within strata to step-one or step-two in the cRCT-SW design. Randomisation was performed offsite by an independent researcher during the third quarter of 2014. As the intervention involves training, specialist mental health care staff knew ~~were aware of~~ their allocated condition as the study progressed.~~.~~ Consumer participants, however, were not informed if staff at their service received the training and efforts were made to maintain the blindness of research assistants for onsite recruitment and stream-two interviews with consumers.~~.~~ Further details are in the protocol paper.[19] ~~[12]~~

Formatted: Right: -1.15 cm

~~Consumer recruitment~~

~~**Consumers were eligible for recruitment if they were: receiving care from a participating cluster with contact in the three months prior to data collection; aged 18-75; able to provide informed consent; and proficient in English. People in prison were excluded. Eligibility was established via a screening process conducted by administration and clinical staff at the participating organisations using detailed instructions provided by the research team. A letter sent to all eligible consumers from each cluster site invited completion and return of a survey form comprising demographic information and the QPR; consent was given by return of the completed survey. An AUD$10 shopping voucher was sent to participants for returned surveys where contact details were provided.**~~ Procedures

~~A range of complementary recruitment strategies to promote consumer response to the mailouts were added according to site need.[12] Strategies included, for example, having researchers, including consumer researchers, speak about PULSAR at participating sites; use of publicity materials such as posters or PULSAR-branded materials; and direct contact with clinicians and consumers at participating sites. Through an active quality assurance process monitoring recruitment, and because this was important to the design, the balance of recruitment approaches was kept as consistent as possible across time points and clusters.~~

Intervention

The REFOCUS intervention[2,18]~~intervention[2,11]~~ introduced earlier as~~was~~ developed in the UK to promote ROP ~~with a basic structure including recovery-promoting relationships and three working practices listed above.~~

~~The essence of the REFOCUS intervention~~ is described in essence in a freely available manual.[18]~~[11]~~

The REFOCUS-PULSAR intervention comprises a manual[20]~~manual~~ adapted from REFOCUS, a structured training intervention to support use of the REFOCUS-PULSAR manual, and follow-up sessions called PULSAR Active Learning Sessions (PALS).

REFOCUS-PULSAR development, following Medical Research Council Guidelines for Complex Interventions,[21]~~[18]~~ and the plan-do-study-act (PDSA) model as a method for controlling and improving process[17]~~process[19]~~ was guided by discussions with the REFOCUS research team, consideration by a Lived Experience Advisory Panel (LEAP), and information from qualitative analysis of group sessions with staff from participating organisations. The content of the REFOCUS manual was substantially retained in the REFOCUS-PULSAR manual[20]~~manual[20]~~ with some amendments to contextualise it for the PULSAR study setting including legal and policy contexts.

Formatted: Right: -1.15 cm

Additions – being less than 25% of the manual - included material related to relapse-signatures and relapse-drills, and material on the CHIME ROP conceptual framework 'Connectedness, Hope, Identity, Meaning, and Empowerment'[12] Empowerment'[6] which was developed during the course of the REFOCUS study. In summary, the REFOCUS-PULSAR intervention was grounded in experience and learning from REFOCUS, research evidence, government policy and law.

The REFOCUS-PULSAR training as developed for step-one occupied two working days, either as a two-day spaced program, or as four half-day spaced sessions, including content and exercises around Recovery-Promoting Relationships and Working Practices; delivery was supported by slide-presentations, a manual PowerPoint slides, manuals, session plans and videos. In a change from the REFOCUS intervention, the training was co-facilitated throughout by professional staff professionals and trainers people with lived experience of mental health problems, including the project's consumer researcher. This which, based on local consultations, was expected to enhance the recovery-orientation of the training. Carer input featured in specific sessions. Quality assurance is described in Appendix 2.

The step-one intervention for clinical services was designed as a two-day session, with the community services training planned as a separate two-day session during the same week. In addition to having two project-employed consumer trainers, trainers were accessed from clinical services for clinical sessions and from the community sector for community sessions. This enabled the inclusion of specialist skills and experience in training delivery.[19] Step-two training was modified based on analyses of participant and trainer evaluations from step-one. Details of changes can be found in Appendix 2. was included in specific sessions. Informed by step-one evaluations, PDSA-based modifications to the second round of training at step-two included earlier introduction of the manual and some reordering so that essential material was covered in the first day of training, with more in-depth exploration and practice on day two. Considerable attention was given in step-two to the interaction between the professional and lived experience co-trainers, seeking to model the interactive style being encouraged for ROP. PALS, offered monthly as hour-long sessions to staff and managers of involved teams to support practice-based implementation of ROP, were facilitated by PULSAR investigators and local trainers.

-

The Control control condition

Standard, standard treatment as delivered through the range of teams introduced above, was governed by national standards,[22] [21] adherence to which is maintained by regular accreditation.

Formatted: Right: -1.15 cm

Consumers of the service often will have their locus of care change in response to changing needs between the more intensive community teams (CATTS, MSTS), residential options including the PARCs, or less intensive community options. Case management in community clinics often functions to coordinate transitions through these levels of care and seeks to ensure that needs for medication, monitoring, supportive, and psychosocial interventions are met. Teams typically have multidisciplinary representation from mental health care disciplines with nursing as the largest single workforce component.

Adverse events

Anticipated possible study-related adverse events included: 1) risk of distress by a participant during an interview; 2) issues related to disclosure of potential self-harm or harm to others 3) risk of harm to staff. A risk-prevention and management protocol was approved by the governing HREC. Participants were provided with written contact details of the manager of the governing HREC for complaints. We did not systematically collect other adverse event information from consumers. For further details of adverse events and complaints procedures see Appendix 3.

Outcomes

Measures

These are divided into "Outcome Measures" (OMs), assessing clinical and personal recovery, and "Experience Measures" (EMs), assessing consumers' experience of health care.

~~Primary outcome measure~~

The primary stream-one OM (see Table 4, protocol paper[19])~~outcome measure (see Table 4, protocol paper[12])~~ was the QPR, a 22-item consumer-rated questionnaire used to assess ~~experience of~~ personal recovery with each item being rated on a 5-point Likert scale ranging from 0 (disagree strongly) to 4 (agree strongly) and higher score indicating increased recovery.[15] While a~~.~~[10] ~~Exploratory factor analysis by the QPR developers suggesting a two-factor structure has not been supported by later studies while a one-factor~~ 15-item scale has been suggested as a ~~briefer and~~ perhaps more robust alternative, this has not been independently validated other than within the 22 item questionnaire.[23] Having ~~.~~[22,23] ~~In this study, Cronbach's alpha was 0·95 for both versions. Since we~~ collected the 22-item version and powered the study based on known properties of this, we retain consistency with our protocol paper and focus on the ~~full scale~~ 22 item score. In this study, Cronbach's alpha was 0·95 for both versions.

Secondary measures in stream-two, both consumer-rated were:

Formatted: Right: -1.15 cm

- EM: The importance of services in recovery questionnaire (INSPIRE) assessing recovery support from a worker[24] has sub-scales of support (20 items) and relationship with worker (7 items) scored by converting the mean of 5-point Likert ratings to a percentage.[24]
- OM: The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) assessing emotional and functional well-being has 14 Likert-scaled items with higher scores indicating greater mental well-being.[19]

Additional measures administered to consumers in stream-two (grouped as OMs and EMs) and reported here include:

EMs

- The Perceived Need for Care Questionnaire (PNCQ) assesses perceptions of mental health care, classifying consumer-identified perceived needs as unmet, partially met or met.[25]
- The Client Satisfaction Questionnaire (CSQ) assesses satisfaction with services.[26]

The Mind Australia Satisfaction Survey (MASS) rates satisfaction with services, staff-consumer service delivery partnerships, and individual service-use outcomes.[27]

~~Sample size~~

~~Primary QPR analysis requires recruitment of 252 separately-sampled consumers at each of T0, T1 and T2, from 14 clusters (see Table 1) with 18 from each cluster and intervention in step-one clusters between T0 and T1 and step-two clusters between T1 and T2, to detect a medium effect-size representing a change in mean score by 6·34.[12] Sample size calculations were based on an intra-cluster correlation coefficient (ICC) of 0·05; significance level 0·05; power 0·80; and published data about distribution properties of the QPR: mean = 46 ± 16. Calculations were done using Stata stepped-wedge V.11.[24]~~

~~Data analysis~~

~~The effect of REFOCUS-PULSAR in the cRCT-SW on the QPR-22 was examined using a 2-level linear-mixed-effects model with robust standard errors. Following 'Intention-to-Treat' analysis requirements, data was grouped by original assigned group[25]. Preliminary models investigating cluster-effect on QPR as random or fixed indicated (likelihood-ratio test p<0·0001) that specifying cluster as random produced better model-fit[26] so in further models, clusters were specified as random. The initial variance-covariance matrix for the random effect was set to unstructured with cluster specified as a random effect while time, intervention status, sector and step-groups were specified as fixed. A simple model with two variables, time and intervention status, and examining any~~

Formatted: Right: -1.15 cm

~~interactions (model 1) was followed by models including covariates of age and gender (model 2), plus sector (model 3), finally adding step-group (model 4). Model 3 examined interactions between three variables (sector, time and intervention), and model 4 examined interactions between four variables (sector, step, time and intervention). After each regression model, a Wald test for significant interaction, with null hypothesis that all interaction coefficients in the model are equal to zero, was conducted. Model adjusted QPR means and the raw data sample QPR means are presented, and broken down by time points, intervention status, sector, and step groups. Model fit was examined using AIC values. Analyses were conducted using Stata V.14.~~

-

- The Coercion Ladder, a visual analogue scale, measures consumers' perception of coercion in mental health service interactions.[28]

OMs

- The Global Assessment of Functioning Scale (GAF) is a researcher-rated (0-100) positively rated measure of individual social, occupational and psychological functioning.[29]
- The Social and Occupational Functioning Assessment Scale (SOFAS), researcher-rated, (0-100) measures function independently from psychological condition severity.[29]
- Days out of role. This measures the impact of mental health problems on usual daily activities over the previous 30 days.

Participant demographic information was also collected.

Consent and key data collection timepoints

In stream-one, consent was by return of a completed survey. Stream-two participants provided written informed consent; interviews took around 60-90 minutes - interviewer blindness was assessed at completion (see protocol paper[19] for further details).

Baseline (T0) data collection occurred in the year prior to and three months after the delivery of the step-one intervention. The first three months after intervention delivery is deemed suitable for baseline data collection based on the Kirkpatrick training evaluation model,[11] whereby the embedding of practice change is considered to take at least 9 months: 3 months for consolidation and 6 months for implementation. During both T1 and T2 periods, data collection at clusters sites took place at a minimum of 9 months after delivery of the intervention to allow embedding of intervention principles and practices.[19]

Formatted: Right: -1.15 cm

Staff finishing REFOCUS-PULSAR training were asked to complete a training evaluation (K1[11]) rating satisfaction from 1 "extremely dissatisfied" to 10 "extremely satisfied". Team managers or administrators were asked to record staff movements every three months.[19] The percentage of the team that attended at least one training session, in both headcount and full-time equivalent (FTE), was calculated for time of training. Team staff turnover was the percentage of staff who left, joined, or moved internally in the organisation but out of the cluster calculated on headcount.

Statistical analysis

Power

These calculations, using the sample size and power calculations described by Hemming and Girling[30] via Stata stepped-wedge V.11[31] were based on: 14 clusters; an intra-cluster correlation coefficient (ICC) of 0·05; significance level 0·05; power 0·80; and published standard deviations.[19] Stream-one and stream-two were powered for medium primary-outcome (QPR) effects. Stream-one detection of a change in mean QPR score by 6·34 indicated 756 surveys (252 in each wave, 18 per cluster per wave). Stream-two detection of a change in mean QPR score by 7.68, indicated 252 surveys (63 at baseline, 126 at step-one and 63 at step-two, 9 per relevant cluster per step). For stream-two secondary outcomes, expected detection thresholds were mean changes in WEMWBS of 4.8 and INSPIRE of 7.72 (medium effects).

Analysis plan

Intention-to-treat analysis was performed in line with a pre-specified analysis plan for all outcomes, using Stata (version 15). Participants were analysed in the groups to which their participating clusters were allocated. We analysed all outcomes using multi-level regression models (linear or Poisson regression as appropriate), with timepoint and intervention status as fixed effects, and clusters as a random effect. Timepoint was included as a categorical variable. Covariates, selected on statistical and clinical considerations, were age-group, gender, sector (PMHS/MHCSS) and step group (stream-one models only). No other covariates have yet been investigated for inclusion into the models, and a later separate investigation will explore the large pool of covariates and their effects on the study outcomes. Covariates of age-group and gender were included as they commonly influence clinical outcomes. Sector (PMHS/MHCSS) was included, as the most important stratification variable, but not the other seven strata as this would have produced an overfitted model. Stream-one models included step group (step-one or step-two) - important temporal changes in the setting and changes in the intervention between steps are detailed in supporting materials. Step group could not be included in stream-two models due to collinearity with intervention status in the incomplete cRCT-SW design.

Formatted: Right: -1.15 cm

It was anticipated (see protocol paper,[19]) that consumers would be modelled as random to account for repeated measures, but stream-one and stream-two cross-sectional recruitment attracted predominantly singletons, contributing to one timepoint only. Simulation studies have found low levels of bias for models with up to 70% singletons and 50 to 500 clustering units[32] so an adjustment to the analysis plan specified that participants would be specified as random only if less than 70% of data came from singletons.

Intervention effects are estimated from the models described above, recommended by Hussey and Hughes.[33] Also investigated and supplied as supplementary analyses in appendices are models with interaction effects between timepoint and intervention status, in which trends across the defined sector (PMHS and MHCSS) are reported.[34] The statistician was not blind to treatment allocation during the analyses.

Role of funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit it for publication.

Ethics approval

Approval was obtained from Monash Health (14102B) and Monash University (CF14/1600 – 2014000773) Human Research Ethics Committees.

Results

Implementation

Training and PALS

Step-one ~~Core~~ REFOCUS-PULSAR ROP training was delivered to 84 staff from the three services in the first quarter of 2015, in 22 days of workshops delivered by 7 trainers. ~~-~~Step-two training was held~~delivery,~~ in June-July (plus an extra session in October) 2016 and~~was~~ delivered to 106 staff over 21 days by 8 trainers. In total 190 staff (111 PMHS; 79 MHCSS) were trained. On average across clusters, 49·1% (PMHS: 38·2%; MHCSS: 63·8%) of staff employed at the time of training attended at least one training session. Adjusted for FTE, this was 51·2% (PMHS: 38·8%; MHCSS: 62·4%). Staff turnover was 42·2% for PMHS, 46·7% for MHCSS-1 and 26·7% for MHCSS-2. ~~Positively judged satisfaction improved significantly from Year 1 to Year 2, Odds Ratio 2·71 (95% CI: 1·04, 7·05, *p* =~~

Formatted: Right: -1.15 cm

~~0·04). PALS sessions had variable uptake, from some teams where implementation was unsuccessful at any level to others where the most scheduled sessions took place and were well attended. Reflective sessions for managers proved not well attended and after a year was ceased in preference for direct reflective discussions carried out by the PALS facilitators.~~

Training satisfaction (K1; scores >5) improved significantly from Year 1 to Year 2, Odds Ratio 2·71 (95% CI: 1·04, 7·05, p = 0·04). Staff trained included representatives of multiple disciplines but the team-based training approach in the most part did not succeed in engaging senior medical staff – it became apparent through the project that they more typically attend service-wide profession-specific trainings which would not readily be compatible with the cRCT model. A medical-specific training of 2 x 1·5 hour sessions was attended by 11 registrars but no consultants.

For two PMHS teams no PALS occurred for logistical and engagement reasons. For all PMHS teams where they did occur (seven team settings including some that were combined), the mean total number of sessions was 8·1, SD 4·7. For 22% of these sessions, arranging team sessions was not successful so meetings were with individual clinicians. In MHCSS settings PALS came to be integrated into monthly staff support sessions and so the element of this that was PALS-specific cannot be quantified.

Consumer Recruitment

Between 18 September 2014 and 19~~31~~ May 2017, 942 consumer participants were recruited across the three time-points, 575 from PMHS and 367 from MHCSS. Of these, 273 participants were recruited for stream-two interviews at timepoints related to the intervention delivery (baseline: 140, follow-up: 133). Overall recruitment targets were surpassed at each time-point (T0, T1 and T2) and most clusters were recruited into as planned (N=18 per cluster) at each time-point (see Figure 1, Figure 2, Appendix 4 and Table 1). As expected, overall recruitment rate from mailouts was low at 8·1%~~%,~~ but yielding 622 or 66% of QPRs. Overall onsite recruitment rate as a proportion of all participants was 39·9% yielding 320 (34%) of all QPRs. Percentages~~As introduced above, consistency of recruitment strategies across time-points was important to successful implementation of the research design and percentages~~ of QPRs derived from onsite recruitment were 32% at T0, 34% at T1 and 36% at T2. Table 1~~Table 1~~ describes each cluster including: organisation sector, stratification level, allocated intervention step, and number of consumer participants recruited at each timepoint in both stream-one and stream-two.~~time point.~~ Table 2 shows the consumer descriptions in stream-one with further details in Appendix 5 which also included details of consumers recruited into stream-two~~the three cross-sectional surveys~~.

Formatted: Right: -1.15 cm

Blindness

This was systematically assessed – see Appendix 6. We see it as unlikely through the course of the project that unblinding represented a significant bias to findings.

Outcomes

Model specification

In line with the adapted analysis plan, since in stream-one and stream-two 90% of the data arose from >50 (854, 254) singletons, consumer was not specified as random.

Primary outcome

The main~~Mixed-effects~~ model outputs in Table 3~~Table 3~~ show that, after adjusting for age~~, and~~ gender and step group and accounting for clustering, we find significant intervention and sector effects. The processes done to build the main model are in Appendix 7 and for the interaction term model in Appendix 8. Figure 3 presents the ~~Wald test results indicate significant interactions for model 1 (p=0·01), model 2 (p=0·02), model 3 (p<0·01), and model 4 (p<0·0001). AIC values indicated that model 2 provided the better fit. ICC for clusters was 0·046. Reference category, having the lowest QPR mean, is the no-intervention group at T1.~~

~~Table 4 shows model-~~adjusted primary outcome means determined by the interaction term model~~for the QPR at each time-point and also raw data sample means. Overall, modelling showed significant intervention effects (p<0·05) reflected in pooled data by the non-intervention clusters mean decreasing 3-4 points at year 1 compared to baseline, and then improving back to baseline levels in year 2 once all clusters had received the intervention. Clusters receiving the intervention in year 1 did not exhibit this decrease in mean. While sector sub-group analyses showed PMHS findings were similar to the overall result, the MHCSS mean remained low in year 2~~.

Table 4 shows intervention effects, estimated as the difference in model-adjusted means (Table 3) between control and intervention data. This was 3·7 (95% Confidence interval: 0.5 – 6.8) for the primary outcome in stream-one, which was significantly greater than zero. To illustrate the degree of the effect size, and while there are some complexities in interpreting this in the context of the specific modelling, we have estimated Cohen's d for the intervention effect as the model adjusted difference (3·7) divided by the sample standard deviation (16·2) = 0·23 , which is a small effect. Appendix 8 shows the model when including interaction terms, and show the overall mean difference between treatment and control groups at year 1 (model 1.6) was 3.7 (95% Confidence interval: 0.6 – 6.8) which was significantly greater than zero.

Figure 3

Formatted: Right: -1.15 cm

~~Figure 2~~ shows QPR scores over time by sector. Pre/post intervention differences occur between T0 and T1 for step-one clusters, and between T1 and T2 for step-two clusters. Therefore, four pre/post intervention scenarios are depicted in this figure (two in each sector). Two of these showed evidence of a significant pre/post intervention difference in QPR scores: in the PMHS sector (2a), in the step-two group there was a significant difference between T1 and T2 of 4·9 (z-score=3·0, p=0·003; Cohen's d estimate = 0.30, small-to-medium effect); and in the MHCSS sector (2b), in the step-one group there was a significant difference between T0 and T1 of 1·1 (z-score=2·7, p=0·006, Cohen's d = 0.07, small effect).

Secondary and other outcomes

Ten sets of results from stream-two are shown in Table 1. Analysis of findings from the PNCQ and a conclusion regarding direction of change are presented in Appendix 9. While none of the findings in Table 5 are individually statistically significant, for nine of ten analyses, central estimates suggested a mean change in the direction favouring the intervention, with estimated effect below the level of change for which the study was powered. If the intervention had no effect, then the probability of each result having direction favouring the intervention is 0·5 and the binomial probability that this would occur nine times from ten results is 0·0107. So the findings suggest some modest positive influences across the span of these variables.

~~In Figure 2, lines are plotted to show the change in non-intervention groups that occurred over the 12 month period between T0 and T1 with a hypothetical (dashed) extension of this forward to T2 in each case. Formal estimation of this trend would carry wide confidence intervals. However viewed in the context of the challenges facing services described earlier, these findings indicate an underlying trend of decreasing QPR scores in the absence of the intervention (reflecting organisational changes impacting on the profile of consumers accessing services), which is perhaps mitigated by receiving the intervention.~~

Discussion

Summary and interpretation of key findings

The ~~PULSAR~~ project found a small but statistically significant effect on consumer stream-one QPR scores for the REFOCUS-PULSAR staff training intervention, ~~using the REFOCUS materials adapted to an Australian context,~~ involving two service sectors and delivered in context of a stepped-wedge design. Small effects in pragmatic trials are expected, and the significant finding is encouraging [35]. A significant interaction effect found for service sector suggests ~~that the~~ changes

Formatted: Right: -1.15 cm

infound in PMHS and MHCSS sectors are better considered separately:. In PMHS, while there was no significant change from T0 to T1 for the step-one group - when this, which might have been expected because this was an they received the REFOCUS-PULSAR intervention period. For the step-two groupduring this time, there was a significant improvement from T1 to T2 (4.92·7 point increase in QPR scores) in theirfor the step-two group, through the time they received the REFOCUS-PULSAR intervention period. In MHCSS, there was small but significant change (1·1) in step-one clusters through their intervention period (T0-T1) and a positive, trend though not significant, trend in step-two clusters from T1 to T2. when their training was delivered.

The 3

A 2·7 point improvement in QPR score represents a 5.73% change in the full scale score.

RecommendationsThis unstandardized metric is the only effect size reported here, as recommendations regarding the modelling approach used are that standardized effect sizes are easily distorted by factors unrelated to size of effect[36]effect[27] and are not straightforward to interpret due to expected variance differences in the mixed model components.[37] Nevertheless the indicative calculation given of Cohen's d suggests a small positive effect so this is how we have framed our discussion. Based on QPR questionnaire content, changes of 1-2 points might be clinically meaningful.[28] Based on QPR questionnaire content, even changes of 1-2 points may be clinically meaningful. For instance a 2 point shift is achieved if the item 'I feel part of society rather than isolated' goes from neutral to strongly agree, which might represent a significant recovery outcome. The training team, working in a PDSA approach, made modifications to the training as delivered in step-two following feedback from step-one. These results seem to confirm that these modifications achieveddid achieve an enhanced impact in PMHS step-two and that the REFOCUS-PULSAR intervention was associated with positive changes in QPR mean scores, most especially as refined for step-two and in the PMHS sector.

While speculative, mechanisms that might have led to greater primary outcome effect in step-two in PMHS might be that the attention given to the relationship between the two trainers (see Appendix 3) had the intended effect of providing better modelling of behaviour for participants through more clearly demonstrating respect for a lived experience perspective and more advanced communication skills. This perhaps also with introduction of dedicated content on coaching. Earlier availability of the manual may have improved uptake of principles for some participants while the team may also generally have gained experience with the delivery of both the core training and the PALS through time. MHCSS findings may be influenced by pressures building in that sector through the course of the project as noted earlier and particularly potentially negatively influencing step-two findings.

Formatted: Right: -1.15 cm

Stream-two findings included non-significant small effects, typically below the study power threshold. While conclusions here must be qualified, in nine of ten instrument comparisons the direction of central estimate of effect was in the direction favouring the intervention condition, a finding unlikely to be due to chance. At least it seems unlikely that any improvements in ROP came at a systematic cost in terms of other impacts. On balance of probabilities it is more likely that there was some small level of clinical and other benefit from the intervention.

Comparisons with REFOCUS

~~This study differed somewhat in design, setting and intervention from the REFOCUS work, while remaining closely related to it.~~ Findings here are more positive overall than those from the REFOCUS study. The differences developed between PULSAR and REFOCUS including those based on learnings from the REFOCUS experience may ~~all~~ have influenced this. The literature on stepped-wedge designs had advanced in the period between design of REFOCUS and PULSAR and the adaptive nature of the PULSAR design allowed for refinements of the training following the first implementation to be evaluated. We note that if this study had been conducted with a similar parallel-group RCT design to that of REFOCUS, then without the inclusion of the step-two findings, PULSAR would not have yielded the positive findings reported here. The involvement of facilitators with lived experience of mental health issues and recovering is central to challenging conventional practices, and in making progress toward an effective recovery-oriented mental health workforce.[7] This might be why we achieved significant finding particularly in step-two PMHS, when the interaction between co-facilitators had been further developed.

Limitations

Accuracy of change-estimates might have been affected by the ~~critical~~ challenges facing the services as noted in the introduction. In both sectors the trend from T0 to T1 in the step-two group receiving no intervention in this time was of declining QPR scores, this most strongly in the MHCSS. Taking into account the challenging influences on all involved services, particularly MHCSS as noted in the 'Setting' section earlier and Appendix 1, it may be that these were acting across the services to drive QPR scores down. If that effect were also operating in the teams at the time they were receiving the intervention, then the underlying trend there might have been towards declining QPR as well. In this case, the findings might be underestimating the effect of the REFOCUS-PULSAR intervention. ~~For instance, in PMHS for step-two, extrapolation of the central estimate of this underlying trend (Figure 2a) would suggest there might have been a positive change of seven QPR points.~~

Formatted: Right: -1.15 cm

REFOCUS-PULSAR training only managed to reach half of staff in intervention sites and few medical staff, which may have reduced intervention potency. In implementation outside constraints of a team-randomised cRCT, better results might be expected from greater engagement of medical staff whether in team-based or profession-specific training.

The REFOCUS intervention recommends some record-structure changes to support ROP, not possible in this cRCT because of organisation-wide regulation of form structures. In the PMHS since PULSAR concluded, the CHIME framework[14] has been integrated as a prompt into an organisation-wide record suite revision which has contributed to further interest in REFOCUS-PULSAR training.

~~Our results are restricted to the QPR because only this data along with demographic details were collected in this main trial, which has the advantage of avoiding multiple analyses. As recommended by CONSORT, in this pragmatic trial the QPR was chosen to be most relevant to the consumers and key decision-makers to whom the trial was aimed.[29] Further information will emerge in future papers when it will be possible to link these findings with results from the other studies included in the PULSAR work program.[12] For example, examination of differential adoption levels in teams may mean the findings are shown to be inconsistent between teams as in REFOCUS;[7] further analyses under way will examine this in the context of this paper's overall positive findings.~~

Our recruitment strategy including~~of~~ repeated sampling and direct consumer approaches was chosen for ~~the~~ strengths of avoiding clinician discretion as a key action-point for selection bias, enhancing consumer autonomy in participation,[38] ~~30~~ and of avoiding selection bias towards greater chronicity of course of illness, identified as a problem in REFOCUS. However, while we have documented the considerable efforts made~~gone to~~ towards consistency of recruitment strategies, the possibility that this created time-variant selection bias on findings cannot be entirely excluded. ~~excluded. The circumstances of the organisations and the findings noted above would suggest that in absence of intervention QPR scores in the consumer sample frame might have been typically going down and in that case, the likely influence of these changes would be to reduce the size of observed effects, not increase them.~~

Formatted: Right: -1.15 cm

Further work

Policy on ROP has been described as "substantially ahead of research and practice"[6]; this is a values-based movement and policy and societal imperatives are strong that something be done to encourage services to work towards ROP even while evidence as to what is best to do may be accumulating. Multiple other ROP based trainings are in use with limited evidence at K1-3 and typically none at K4. The REFOCUS-PULSAR program can be considered for use based on reported findings suggesting improvements in high adopting teams in the English study, along with these K4 findings from PULSAR. There~~Specific service recommendations~~

~~Recommendations made from the study team to PMHS and MHCSS service partners~~ have been requests from teams in~~to rerun~~ the participating PMHS for further ~~cycle of REFOCUS~~ PULSAR training, with exploration of ~~three times in the next year,~~ extending and adapting the training to include inpatient staff so that the recovery-oriented culture can extend more widely across the care spectrum. In responding to these requests this team are mindful of the need to continue carefully to evaluate such initiatives, continuing PDSA cycles also with attention to educational evaluation at levels K1-K4 wherever possible and development of fidelity measures.

To better understand how sustained practice change can be achieved within services, future ROP training initiatives are recommended to strengthen~~; and work to increase~~ the focus on implementation strategies, such as follow up coaching or mentoring, refresher programs, and service user feedback and evaluation.[8,10,39] Wide-ranging organisational factors are recognized as influential in supporting or constraining ROP implementation efforts,[9,10] so that attention to organisational readiness for change and alignment of organisational policies, processes, staffing and resources with recovery oriented principles are also important. cRCT designs studying teams impede use of ~~integration of REFOCUS-PULSAR documents with~~ organisation-wide strategies and RCTs where randomisation is by organisation~~-record systems and forms. These~~ have limitations of large clusters so design considerations continue to be a challenge in accumulation of highest level evidence for these approaches. ~~all been accepted by participating organisations and prompts related to the REFOCUS-PULSAR frameworks are now integrated into the PMHS treatment and recovery plan documentation.~~

Conclusions

Taken together, these results suggest that the REFOCUS-PULSAR intervention can lead to a modest overall measured improvement in personal recovery, also possibly with a small effect on some measures of clinical recovery and other aspects of client experience.[40] From an educational

Formatted: Right: -1.15 cm

intervention perspective they place the REFOCUS-PULSAR intervention in the situation of having at least some evidence at level K4,[11] something otherwise lacking in the literature surveyed to date. It seems at very least unlikely that any improvement in ROP came at a cost in terms of clinical measures. While the findings of this study are modest, this is not surprising in a pragmatic trial and they provide at least some indication of positive change for consumers accessing the intervention services.

~~This intervention was effective in promoting ROP across sectors, though this effect was only clearly seen in Public Mental Health Services after the training had been refined. The intervention should be considered for services as part of broader organisation-wide initiatives to promote recovery-oriented practice.~~

Contributors

GM was the Principal Investigator on this trial and together with JE led development of key elements of the design and analysis approach and interpretation of the findings. MS developed the original REFOCUS intervention and advised on adaptation. LB chaired the research module task-group and provided oversight to development and implementation of all elements of the design. FS provided overall coordination for field work and staff training and was centrally involved in the day-to-day operations of trial implementation. JE conducted the analyses. Specialist contributions regarding design elements were made by EF and EWE. CT contributed to the study design and implementation within MHCSS and was a chair of the adaptation module task-group. PW chaired the implementation module task-group which oversaw the delivery of the training intervention. VE, LB, GM, PW and EWE developed the specific specialist care training intervention and associated resources. GM, LB, FS, VE and EWE developed the specialist care instrumentation and fieldwork trial protocols. The core drafting group for this paper comprised GM, LB, FS, JE, MS; the remaining authors revised it critically for important intellectual content. All authors read and gave final approval for this version of the paper to be published.

Declaration of interests

We declare no competing interests.

Formatted: Right: -1.15 cm

Data sharing statement

In compliance with the requirements of the Monash Health Research Ethics Committee, the data supporting our findings in the manuscript will not be shared because we did not obtain participant consent to do so.

Acknowledgements

We thank: the partner organisations of ermha, Mind Australia and Monash Health for support to the project; the Steering Group and Lived Experience Advisory Panel for sharing their expertise; the REFOCUS-PULSAR trainers; and the team of PULSAR research and fieldwork staff for data collection. We thank all project participants who in sharing their time and experience made this work possible. This article presents independent research funded by the Victorian Government through its Mental Illness Research Fund. MS is supported by the NIHR Nottingham Biomedical Research Centre. The views expressed in this publication are those of the authors and do not necessarily reflect those of the Victorian Government including the Department of Health and Human Services. Further information is available at www.pulsarrecovery.org.au.

Formatted: Right: -1.15 cm

Formatted: Right: -1.15 cm

Tables and Figures

Table 1 consumer numbers by cluster~~Cluster sites~~, stratification levels, intervention step~~,~~ and timepoint

(a) Stream-one trial numbers~~number~~ of consumer participants in the three cross-sectional surveys who completed the Questionnaire about the Process of Recovery (QPR)

	Site information				QPR surveys		
Cluster	**Organisation**	**Strata**	**N**	**%**	**T0**	**T1**	**T2**
1	PMHS	A	66	7·0	23	29	14
2	PMHS	B	37	3·9	14	12	11
3	PMHS	C	66	7·0	21	24	21
4	PMHS	D	104	11·0	32	38	34
5	MHCSS - 1	E	52	5·5	16	15	21
6	MHCSS – 2	F	64	6·8	20	17	17
7	MHCSS - 1	G	56	5·9	19	25	12
8	PMHS	A	98	10·4	30	26	42
9	PMHS	B	44	4·7	17	15	12
10	PMHS	C	89	9·5	21	41	27
11	PMHS	D	71	7·5	20	26	25
12	MHCSS - 1	E	69	7·3	21	24	24
13	MHCSS – 2	F	52	5·5	21	17	14
14	MHCSS – 2	G	74	7·9	26	25	23
		Total, All sectors	***942***	***100***	***301***	***334***	***307***

Notes. Clusters were stratified by the team/service type and composition: i.e. seven different strata. Overall there were 575 (61·0%) consumer QPR surveys from Public Mental Health Services (PMHS) and 367 (39·0%) from Mental Health Community Support Services (MHCSS): 177 (18·8%) from MHCSS-1 and 190 (20·2%) from MHCSS-2.

(b) Stream-two trial numbers of consumer participants who participated in a study interview

Site information					QPR surveys		
Cluster	Organisation	Strata	N	%	T0	T1	T2
1	PMHS	A	22	8·1	10	12	-
2	PMHS	B	15	5·5	10	5	-
3	PMHS	C	17	6·2	6	11	-
4	PMHS	D	24	8·8	14	10	-

5	MHCSS - 1	E	11	4·0	9	2	-
6	MHCSS – 2	F	19	7·0	11	8	-
7	MHCSS - 1	G	23	8·4	11	12	-
8	PMHS	A	26	9·5	-	13	13
9	PMHS	B	7	2·3	-	5	2
10	PMHS	C	29	10·6	-	12	17
11	PMHS	D	16	5·7	-	9	7
12	MHCSS - 1	E	23	8·4	-	9	14
13	MHCSS – 2	F	18	6·6	-	12	6
14	MHCSS – 2	G	23	8·4	-	9	14
	Total, All sectors	***273***	***100***	***71***	***129***	***73***	

Notes. Clusters were stratified by the team/service type and composition: i.e. seven different strata. Overall there were 156 (57·1%) interviews with consumers~~575 (61·0%) consumer QPR surveys~~ from Public Mental Health Services (PMHS) and 117 (42·9~~367 (39·0~~%) from Mental Health Community Support Services (MHCSS): 57 (20·9~~177 (18·8~~%) from MHCSS-1 and 60 (22~~190 (20·2~~%) from MHCSS-2.

Key:
Control condition period
Intervention condition period

Table 2 Stream-one trial QPR numbers (%) by timepoint, gender, age group, step, intervention status and demographics

~~Key:~~	Timepoint
~~Control condition period~~	
~~Intervention condition period~~	

~~Table 2 QPR numbers by time point, gender, age group, step, intervention status and demographics~~

	T0	T1	T2	Total
Distribution in specialist care by Timepoint~~Time Point~~				
N	301	334	307	942
(%)%	(32·0)	(35·5)	(32·6)	(100)
Distribution in specialist care by Timepoint~~Time Point~~ and Gender[1]				
Female	174 (57·8)	192 (57·5)	178 (58·0)	544 (57·7)
Male	125 (41·5)	139 (41·6)	126 (41·0)	390 (41·4)
Not listed	2 (0·7)	3 (0·9)	3 (1·0)	8 (0·8)
Distribution in specialist care by Timepoint~~Time Point~~ and Age group				
17-30 years	73 (24·3)	77 (23·1)	79 (25·7)	229 (24·3)
30-49 years	151 (50·2)	170 (50·9)	151 (49·2)	472 (50·1)
50 years and over	72 (23·9)	84 (25·1)	74 (24·1)	230 (24·4)
Distribution in specialist care by Timepoint~~Time Point~~ and Step Group intervention				
Step Group 1	145 (48·2)	160 (47·9)	140 (45·6)	445 (47·2)
Step Group 2	156 (51·8)	174 (52·1)	167 (54·4)	497 (52·8)
Distribution in specialist care by Timepoint~~Time Point~~ and Intervention status (Ix)				
No Ix	301 (100)	174 (52·1)	0 (0·0)	475 (50·4)
Yes Ix	0 (0)	160 (49·9)	307 (100)	467 (49·6)
Distribution in specialist care by Country of birth				
Australia	217 (72·1)	244 (73·1)	229 (74·6)	690 (73·2)
Other	83 (27·6)	87 (26·0)	73 (23·8)	243 (25·8)
Not listed	1 (0·4)	3 (0·9)	5 (1·6)	9 (1·0)
Distribution in specialist care by Year of arrival				
After 2000	17 (5·6)	23 (6·9)	19 (6·2)	59 (6·3)
Between 1981-2000	40 (13·3)	39 (11·7)	27 (8·8)	106 (11·3)
Before 1980	18 (6·0)	17 (5·1)	17 (5·5)	52 (5·2)
Not listed	8 (2·7)	8 (2·4)	10 (3·3)	26 (2·8)
Distribution in specialist care by Main language				
English	265 (88·0)	286 (85·6)	269 (87·6)	820 (87·0)
Other	23 (7·6)	26 (7·8)	23 (7·5)	72 (7·6)
Both English and Other	8 (2·7)	17 (5·1)	7 (2·3)	32 (3·4)
Not listed	5 (1·7)	5 (1·5)	8 (2·6)	18 (1·9)
Distribution in specialist care by Ethnicity (self-identified)				
Australian Non-Indigenous	121 (40·2)	177 (53·0)	162 (52·8)	460 (48·8)
Australian Indigenous	27 (9·0)	20 (6·0)	33 (10·7)	80 (8·5)
Other	120 (39·9)	126 (37·7)	97 (31·6)	343 (36·4)
Not listed	33 (11·0)	11 (3·3)	15 (4·9)	59 (6·3)
Other category (multiple responses could be listed)				
English, Irish, Walsh, Scottish	25 (8·3)	42 (12·6)	29 (9·4)	96 (10·2)
Italian	13 (4·3)	17 (5·1)	10 (3·3)	40 (4·2)
Greek	7 (2·3)	17 (5·1)	11(3·6)	35 (3·7)
New Zealander/Maori	11 (3·7)	10 (3·0)	12 (3·9)	33 (3·5)
Other (participant selected "other")	72 (23·9)	58 (17·4)	33 (10·7)	163 (17·3)
Censored[2]	108 (35·9)	91 (27·2)	76 (24·8)	275 (29·2)
Distribution in specialist care by Duration of mental health service use				
Mean number of years	4·0	4·5	4·0	4·2
Median number of years	1·0	1·0	1·0	1·0
Range (years)	0-35	0-35	0-35	0-35
No. of people with <1 year at site	129 (42·9)	125 (37·4)	135 (44·0)	389 (42·3)
Mean number of months for those with <1 year at site	3·3	3·2	3·2	3·2
Median number of months for those with <1 year at site	3	3	3	3

Note. Where cell sizes are less than 5 at any timepoint~~time point~~ for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other _____.

[2]Included 56 additional ethnic groups

Table 3 Stream-one QPR mixed model~~Model~~ **coefficients** with~~produced by the 2-level mixed regression models~~ ~~(A) Model 1 has~~ **fixed** factors~~variables~~ **of** sector, step-group, sex, age-group, **time**-point **and intervention status**, with clusters as random. Number of observations=942.

		b	se	z	p value	ll	ul
sex							
	Female	-0·81	1·04	-0·79	·431	-2·84	1·21
Age group							
	2	-0·94	0·88	-1·07	·285	-2·65	0·78
	3	-3·4	0·91	-3·78	0·001***	-5·22	-1·66
Timepoint							
	T1	-3·22	1·02	-3·16	·002***	-5·22	-1·22
~~**Time point**~~	**Intervention** status			~~**Coefficient**~~	~~**Robust Std Err.**~~	~~**P>|z|**~~	~~**95% CI**~~
	T2	-4·22	1·50	-2·82	·005***	-7·15	-1·29
	yes	3·76	1·31	2·87	·004***	1·20	6·33
Sector							
	2	-1·72	2·12	-0·81	·418	-5·87	2·43
Step group							
	2	0·15	2·08	0·07	·943	-3·93	4·22

*** p<0·01, ** p<0·05, * p<0·1

Table 4 Steam-one model-adjusted QPR means - derived from model in Table 3

		QPR mean	Std.Err.	[95%Conf.	Interval]
sex					
	1	54·85	1·36	52·18	57·51
	2	54·03	1·23	51·63	56·44
Age group					
	1	55·69	1·17	53·39	57·99
	2	54·75	1·25	52·31	57·20
	3	52·25	1·48	49·34	55·16
Timepoint					
	0	56·89	1·25	54·45	59·34
	1	53·67	1·43	50·86	56·48
	2	52·67	1·38	49·97	55·37
Intervention status					
	0	52·51	1·46	49·65	55·37
	1	56·27	1·23	53·87	58·67
Sector					
	1	55·05	1·75	51·62	58·47
	2	53·33	1·24	50·91	55·75
Step					
	1	54·29	1·42	51·52	57·07
	2	54·44	1·69	51·12	57·76

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0·5 – 6·8).

Table 5 Summary of outcomes in the streams 1 and 2 trials

Primary outcome:		**Control (n=475)**	**Intervention (n=467)**	**Adjusted diff in means (95%CI); p-value**	**Change in direction of favouring the intervention**
QPR Stream-one	Mean (sd)	53·6 (16·3); n=475	54·4 (16·2); n=467	3·72 [d] (0·51,6·92); 0·023	
Secondary outcomes:		**Control (n=140)**	**Intervention (n=133)**		Yes
QPR Stream-two	Mean (sd)	53·1 (14·8); n=138	54·0 (14·5); n=131	2·54[d] (-3·10,8·18); 0·38	
Warrick	Mean (sd)	41·4 (11·2); n=139	42·2 (11·1); n=133	2·39 [d] (-2·66,7·43); 0·35	Yes
INSPIRE S score	Mean (sd)	62·4 (22·3); n=128	62·2 (23·1); n=123	2·03 [d] (-6·72,10·78); 0·65	Yes
INSPIRE R score	Mean (sd)	72·0 (22·3); n=134	75·5 (20·1); n=129	3·29 [d] (-3·39,9·97); 0·34	
Other outcomes:		**Control (n=140)**	**Intervention (n=133)**		
GAF score	Mean (sd)	48·5 (14·7); n=140	51·4 (13·3); n=133	0·92 [d] (-6·15, 8·00); 0·80	Yes
SOFA score	Mean (sd)	49·8 (15·5); n=134	52·9 (14·3); n=132	0·57 [d] (-5·30, 6·45); 0·85	Yes
Client Satisfaction Questionnaire (CSQ)	Mean (sd)	23·3 (5·3); n=139	24·5 (5·5); n=130	1·21 [d] (-0·98, 3·41); 0·28	Yes
Mind Australia Satisfaction Survey (MASS)	Mean (sd)	8·0 (1·8); n=140	8·2 (1·8); n=132	0·02 [d] (-0·62, 0·67); 0·94	Yes
The Coercion Ladder, Community services	Median (IQR)	2·0 (1·5); n=139	2·0 (1·5); n=139	0·20 [ll] (-1·12, 0·72); 0·67	Yes
Days out of role (full)	Median (IQR)	6·5 (0·0,15·0); n=138	6·0 (0·0,15·0); n=133	-1·37 (-5·34, 2·59); 0·50	No

Days out of role (partial)	Median (IQR)	6·0 (0·0,15·0); n=133	10·0 (2·0,15·0); n=129	0·12 (-4·56, 4·81); 0·96	
PNCQ (see Appendix 7)					Yes

Mean and standard deviation (sd) unless otherwise indicated. Also shown are the adjusted differences calculated from the multi-level mixed models (d linear or II Poisson regressions) adjusted for fixed effects of gender, age, timepoint and sector and clusters as random effects. Step group is an additional fixed effect in the stream-one regressions.

Figures

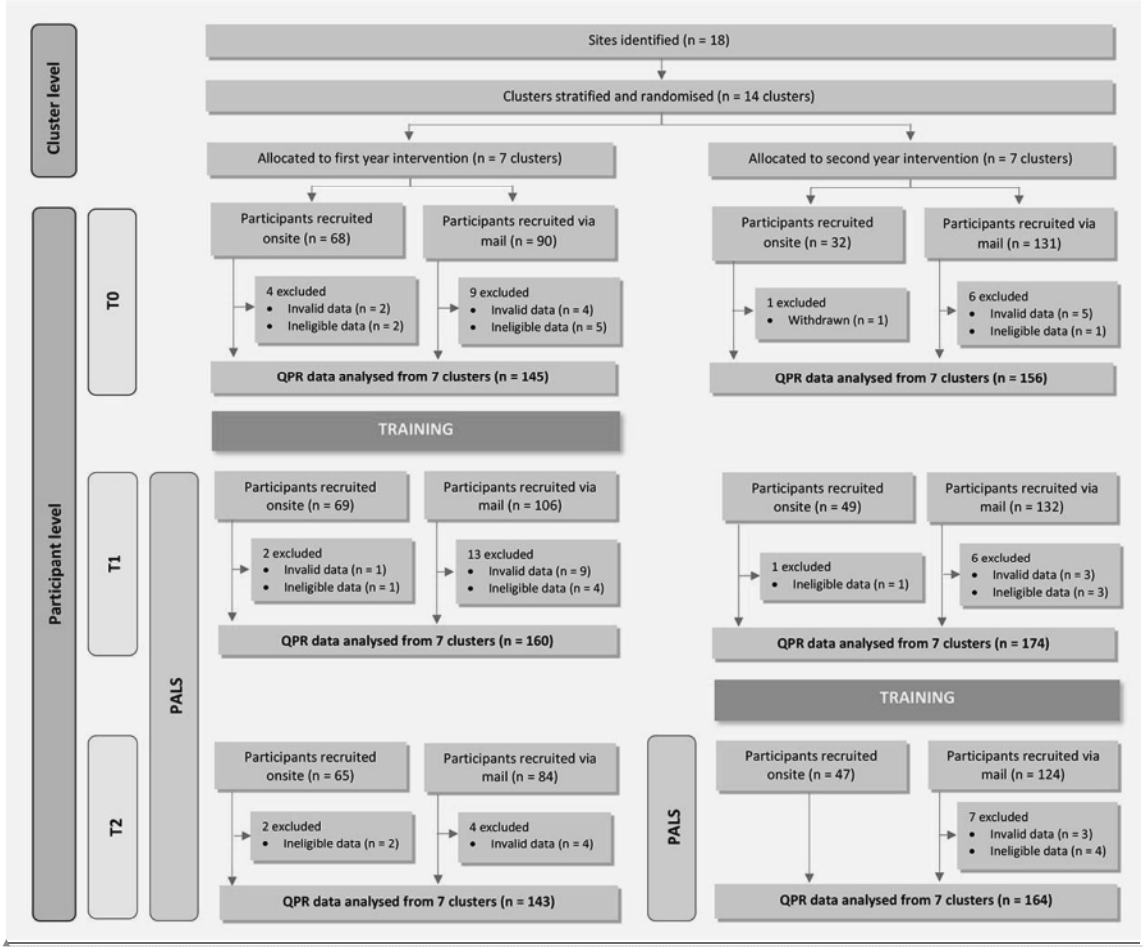


Figure 1 Consort chart for stream-one

Notes. PALS = PULSAR Active Learning Sessions. Invalid data refers to data-based issues in the form of missing data or invalid responses. Ineligible data refers to participant-based issues – that is, the person providing the data did not meet the eligibility criteria for the study.

Formatted: Font: 12 pt

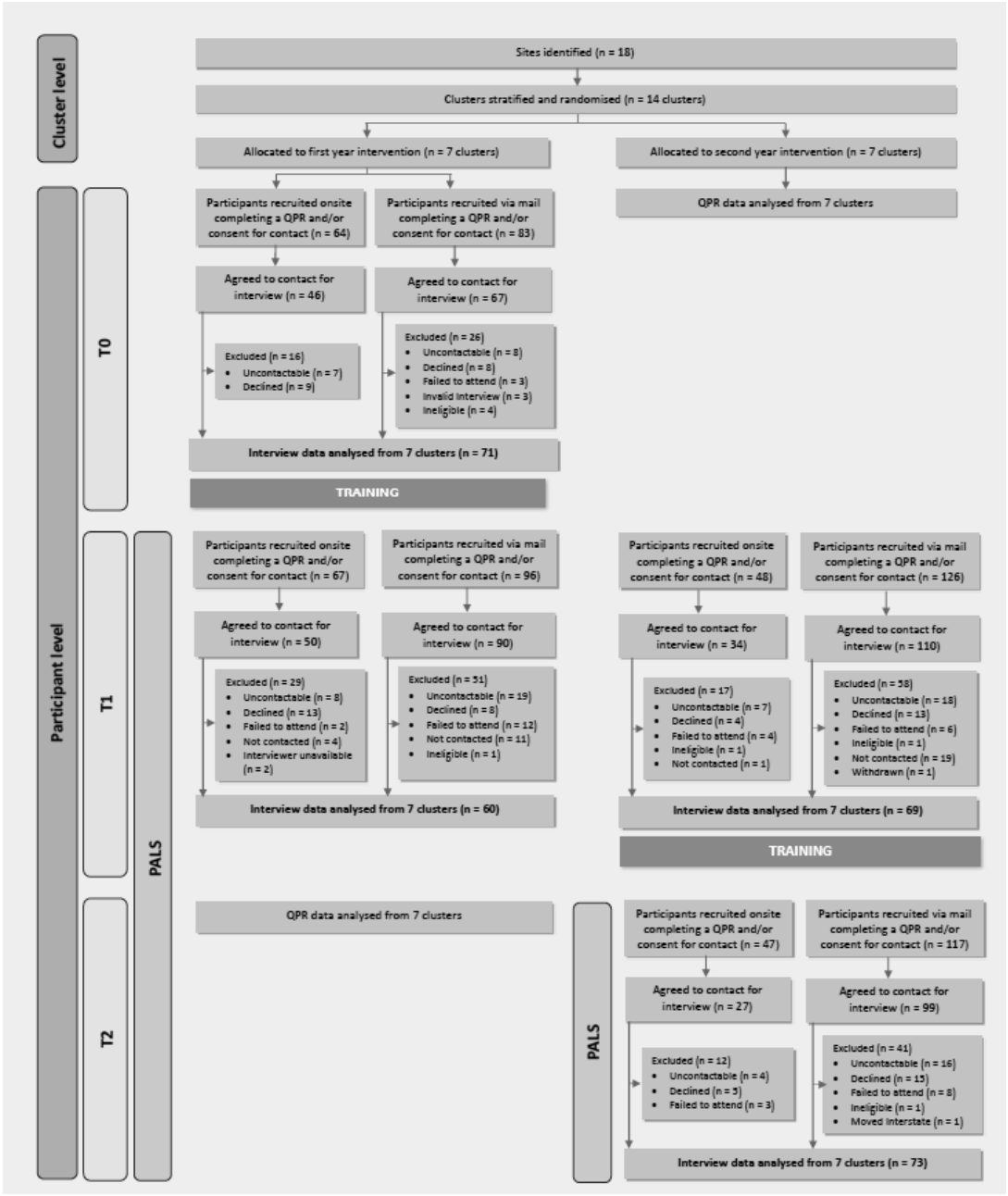


Figure 2 Consort chart for stream-two

Note. PALS = PULSAR Active Learning Sessions

Figure 3a: Model adjusted mean QPR scores over time: Public Mental Health Services (PMHS)

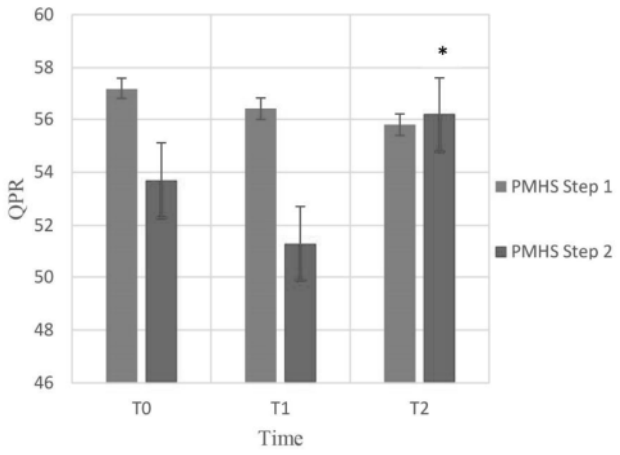


Figure 3b: Model adjusted mean QPR scores over time: Mental Health Community Support Services (MHCSS)

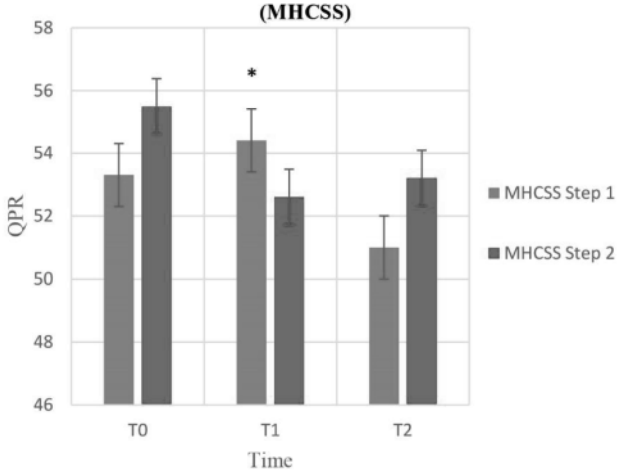


Figure 3 QPR scores by sector over time.

*Change p < .01 by pairwise comparison with previous time-point.
Note. Step-one group (blue) received intervention in year 1. Step-two group (red) received intervention in year 2.

Formatted: Font: 8 pt

References

1. Deegan P. Recovery as a journey of the heart. *Psychiatr Rehabil J* 1996; **19**(3): 91-7.

2. Slade M, Bird V, Le Boutillier C, et al. Development of the REFOCUS intervention to increase mental health team support for personal recovery. *Br J Psychiatry* 2015; **207**(6): 544-50.

3. Shepherd G, Boardman J, Slade M. Making recovery a reality. London: Sainsbury Centre for Mental Health 2008.

4. World Health Organization. Mental health action plan 2013 - 2020. Geneva: WHO Document Production Services; 2013.

5. Australian Government Department of Health. National mental health policy 2008. Canberra: Commonwealth of Australia 2009.

6. Slade M, Bird V, Clarke E, et al. Supporting recovery in patients with psychosis through care by community-based adult mental health teams (REFOCUS): a multisite, cluster, randomised, controlled trial. *Lancet Psychiatry* 2015; **2**(6): 503-14.

7. Salkeld R, Wagstaff C, Tew J. Toward a new way of relating: an evaluation of recovery training delivered jointly to service users and staff. *J Ment Health* 2013; **22**(2): 165-73.

8. Deane FP, Andresen R, Crowe TP, Oades LG, Ciarrochi J, Williams V. A comparison of two coaching approaches to enhance implementation of a recovery-oriented service model. *Adm Policy Ment Health* 2014; **41**(5): 660-7.

9. Uppal S, Oades LG, Crowe TP, Deane FP. Barriers to transfer of collaborative recovery training into Australian mental health services: implications for the development of evidence-based services. *J Eval Clin Pract* 2010; **16**(3): 451-5.

10. Lodge AC, Kaufman L, Stevens Manser S. Barriers to Implementing Person-Centered Recovery Planning in Public Mental Health Organizations in Texas: Results from Nine Focus Groups. *Adm Policy Ment Health* 2017; **44**(3): 413-29.

11. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs. The Four Levels. Third ed. San Franciso, California: Berrett-Koehler Publishers; 2006.

12. Slade M, Bird V, Le Boutillier C, Williams J, McCrone P, Leamy M. REFOCUS Trial: protocol for a cluster randomised controlled trial of a pro-recovery intervention within community based mental health teams. *BMC Psychiatry* 2011; **11**(1): 185.

13. Ajzen I. The theory of planned behaviour: reactions and reflections. *Psychol Health* 2011; **26**: 1113–27.

14. Le Boutillier C, Chevalier A, Lawrence V, et al. Staff understanding of recovery-orientated mental health practice: a systematic review and narrative synthesis. *Implement Sci* 2015; **10**(1): 1-14.

15. Neil ST, Kilbride M, Pitt L, et al. The questionnaire about the process of recovery (QPR): A measurement tool developed in collaboration with service users. *Psychosis* 2009; **1**(2): 145-55.

16. Chow SC, Chang M. Adaptive design methods in clinical trials - a review. *Orphanet J Rare Dis* 2008; **3**: 11.

17. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. *BMJ Quali Saf* 2014; **23**: 290-8.

18. Bird V, Leamy M, Le Boutillier C, Williams J, Slade M. REFOCUS: Promoting Recovery in Mental Health Services. 2nd ed. London: Institute of Psychiatry, King's College London; 2014.

19. Shawyer F, Enticott JC, Brophy L, et al. The PULSAR Specialist Care protocol: a stepped-wedge cluster randomized control trial a training intervention for community mental health teams in recovery-oriented practice. *BMC Psychiatry* 2017; **17**: 172.

20. PULSAR-REFOCUS team. PULSAR Manual: Recovery-promoting Relationships and Working Practices for Specialist and Community Mental Health Services (or Secondary Care Staff). 2nd ed. Melbourne: Monash University; 2016.

21. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions an introduction to the new Medical Research Council guidance. In: Kelly MP, Killoran A, eds. Evidence–Based Public Health: Effectiveness and Efficiency. Oxford: Oxford University Press; 2010: 185-203.

22. Australian Government. National Standards for Mental Health Services 2010. Australia: Commonwealth of Australia; 2010.

23. Williams J, Leamy M, Pesola F, Bird V, Le Boutillier C, Slade M. Psychometric evaluation of the Questionnaire about the Process of Recovery (QPR). *Br J Psychiatry* 2015; **207**(6): 551-5.

24. Williams J, Leamy M, Bird V, et al. Development and evaluation of the INSPIRE measure of staff support for personal recovery. *Soc Psychiatry Psychiatr Epidemiol* 2015; **50**(5): 777-86.

25. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). *Soc Psychiatry Psychiatr Epidemiol* 2000; **35**(9): 427-35.

26. Attkisson CC, Zwick R. The Client Satisfaction Questionnaire: psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982; **5**(3): 233-7.

27. Wray M, Brophy L, Loh M. The Prevention and Recovery Care Exit Survey: Mind Australia; 2013.

28. Høyer G, Kjellin L, Engberg M, et al. Paternalism and autonomy: A presentation of a Nordic study on the use of coercion in the mental health care system. *Int J Law Psychiatry* 2002; **25**(2): 93-108.

29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). Washington, DC, USA: American Psychiatric Association;; 2000.

30. Hemming K, Girling A. A menu-driven facility for power and detectable-difference calculations in stepped-wedge cluster-randomized trials. *Stata J* 2014; **14**(2): 363-80.

31. Statacorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.

32. McNeish DM, Stapleton LM. The Effect of Small Sample Size on Two-Level Model Estimates: A Review and Illustration. *Educ Psychol Rev* 2016; **28**(2): 295-314.

33. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007; **28**(2): 182-91.

34. Hemming K, Taljaard M, Frobes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. *Trials* 2017; **18**(101).

35. Patsopoulos N. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosc* 2011; **13**(2): 217-24.

36. Selya AS, Rose JS, Dierker LC, Hedeker D, Mermelstein RJ. A practical guide to calculating Cohen's f(2), a measure of local effect size, from PROC MIXED. *Front Psychol* 2012.

37. Brysbaert M, Stevens M. Power analysis and effect size in mixed effects models: A tutorial. *J Cogn* 2018; **1**(1).

38. Sharkey K, Savulescu J, Aranda S, Schofield P. Clinician gate-keeping in clinical research is not ethically defensible: an analysis. *J Med Ethics* 2010; **36**(6): 363-6.

39. Stuber J, Rocha A, Christian A, Johnson D. Predictors of Recovery-Oriented Competencies Among Mental Health Professionals in One Community Mental Health System. *Community Ment Hlt J* 2014; **50**(8): 909-14.

40. Le Boutillier C, Leamy M, Bird VJ, Davidson L, Williams J, Slade M. What does recovery mean in practice? A qualitative analysis of international recovery-oriented practice guidance. *Psychiatr Serv* 2011; **62**(12): 1470-6.

Supplementary materials

Appendix 1

Key changes in the setting of the project through the time-period for observations

The key period in which this intervention and associated observations across multiple healthcare sectors in Victoria occurred from late 2014 to early 2017 was a time of considerable external change, collective stresses and challenge to involved organisations and their staff. This context is likely to have had some negative influences on implementation of recovery orientated practice.

The funding environment for public health services in Victoria under the Liberal administration 2010 to 2015 received significant criticism as negative, with funding levels not keeping up with inflation, or where they did, being associated with substantial additional commitments so not representing real increases.[1] While the Labor administration that followed has been better reviewed for its support of healthcare,[2] there was limited time for the actions of this new administration to flow through into changes in work context in the timespan of this project.

As well as the general problem of under resourcing, three intersecting areas of change impacted on this project and the research undertaken with specialist mental health services.

Changes to the MHCSS:

In 2015, the then Victorian State Government introduced a major reform of mental health community support services (MHCSS), which presented substantial challenges for organisations involved in this study.[3] In 2014, Kim Koop, the then CEO of the peak body for mental health community support services (MHCSS) in Victoria, anticipated that MHCSS were about to experience "an extended period of uncertainty" (p.10)[4] as a result of new service types and new contracts being implemented to begin to reshape service delivery that would also herald significant changes in the service provider landscape.[4] This process was seen as necessary to deal with a long-standing need for reform and also to assist the sector to prepare for the much larger transition process of having all MHCSS funding rolling into the National Disability Insurance Scheme (NDIS) by 2019. Later referred to as "recommissioning", Kate Silburn wrote, after interviewing senior personnel, that:

> In the case of recommissioning MHCSS and AOD (Alcohol and other drug) sectors in Victoria there would appear to have been very high costs, potentially avoidable difficulties and as yet undetermined benefits. One of strongest themes emerging from the interviews was that while there had been widespread enthusiasm for reform in both sectors, many of its proponents had become disillusioned and despairing as a consequence of the processes for both reform design and reform implementation (p.39).[3]

The recommissioning involved the establishment of new catchment areas, a centralised intake and assessment process, and a reduction in the number of providers. Many small services lost their funding altogether and larger services had both losses and gains that required considerable adjustment. For example some services were allocated to other providers and staff and clients needed to transition from one provider to another. This resulted in a long and difficult transition period. The costs to the sector were considerable. Agencies, including those in this study, had to make staff redundant and also explain to consumers why they were no longer going to provide services to them. Other problems identified by Silburn[3] are:

- Poor planning, lack of guidelines and information and limited systems alliance (p.19)
- Not enough time and too many things happening at once (p.21)
- Lack of communication with consumers and other types of service providers (p.21)
- Lack of a well-planned process for 'transitioning' clients (p.21)

Silburn (2014) also found that the recommissioning process had undermined:

- collaboration, partnerships and joint models of care (p.24);
- models of care for clients from disadvantaged or vulnerable communities (p.25);
- comprehensive models of care delivered by single agencies

There were concerns that the central intake system became a barrier for clients to access services because it was often multi-stepped and difficult to negotiate. For example:

> Interviewees argued that while clients had previously been able to walk into their agency, make an appointment and get an assessment within a short time period, they now had to be directed to call the central intake provider and may have to wait several weeks for an assessment (p.29)[3]

After recommissioning Silburn[3] describes how:

> MHCSS sector clients are categorised into three tiers, consistent with the proposed categorisation for the NDIS. To be eligible for a service clients have to have a permanent disability associated with a mental illness. Once clients are deemed eligible they are then categorised based on the severity of the disability and/or their current needs. This means that clients with high levels of disability, but who are otherwise stable/doing well (and therefore might have low levels of need) might get the same level of priority as someone who has a lower level of disability and a high level of current need. One interviewee noted that this also means that clients with either no permanent disability or with an uncertain diagnosis, but with high needs are likely to miss out on services in the new system. People in this group can include people experiencing their first psychosis or life circumstances like deterioration of their support networks, who with early intervention may not become

dependent on the MHCSS system. The system has therefore lost significant capacity for prevention for this group of clients. (p.32).

Changes as a result of the new Mental Health Act:

In 2014, the Victorian State government introduced a new Mental Health Act.[5] Although the Act is encouraging of a 'recovery' orientation, this is not explicitly defined in the Act, and was introduced along with its further training requirements in the context of this budget-limited environment. The introduction of the new Mental Health Act in 2014 was the first such change in nearly three decades. It provided an entirely new legal framework for the delivery of mental health services (MHS) and carried with it considerable training demands on clinicians in the public MHS involved. This led to some delay in the ability to deliver the PULSAR training since it was not practicable to release staff in those services so quickly following this other major training impost. The MHA was designed to support recovery and introduce a range of new mechanisms to improve patients or consumers involvement in decision making including Advance Statements and Nominated Persons. The immediate impact of the legislative changes was to produce an uncertain legal environment in which the voluntary PULSAR training modules competed with compulsory training on the Act for staff time and attention. Also this compulsory training tended not to focus on the relevant recovery and supported decision making reforms but rather on the changes to compulsory admission criteria and treatment orders, restrictive interventions and ECT.

Changes at the Clinical Services:

Some quantitative summary data gives an indication of the trends in activity through time across the PMHS, based on regularly collected data and available reports which are not available in as standardised a way for the MHCSS. The last day of the year snapshot of all PMHS case managed clients rose from 2349 in 2014 to 2462 in 2016, an increment of 5%. By way of indicators on demand factors for the whole service, emergency department presentations increased from 8803 in 2014 to 10004 in 2016 (+14%) and inpatient length of stay decreased on average from 12·8 days to 11·3 days (-4%) as total in-patient separations increased from 3102 to 3633 (+17%). Average length of stay in community services from opening to closing of administrative cases increased by 31% (2014: 157 days, 2015: 170 days, 2016: 205 days).
From an observational and more qualitative perspective, Monash Health Service (MH) could be described as a hyper-complex environment [6] and it was particularly so during the time this study was being conducted. In 2015, among significant changes within MH through the course of the project, it is publicly available information that the staff employed as MH Medical Program Director and Executive Director left in April and May of that year respectively. Long et al[7] describe, in reporting on a project that was also occurring during this time, how "MHS underwent a major restructuring after a significant number of senior staff left the service" (p.2).
Long, McDermott and Meadows (as yet unpublished PhD research[8]) describe, via semi structured interviews carried out with the MHS senior leadership group, the amount of change and challenges occurring in the service between 2013 and 2017. While their investigation was a different project to PULSAR it occurred in a similar timeframe and the findings are very relevant to describing the context. Long et al's participants used critical reflection to identify meaningful events in the services during this time.[8] Twenty-three critical incidents were identified. These included changes in government policy, adjustments in funding and staff turnover. Hence staff in the service were persistently having to deal with change and also the loss of some programs, creating an atmosphere of uncertainty.

Appendix 2

Training quality assurance and adaptations made to the training in course of the project

Quality assurance measures employed during training delivery included a day-long workshop attended by all trainers to introduce the training schedule, content and process including demonstrations and role-play of key exercises. A detailed schedule guided delivery working through the key elements of content along with use of standardised training materials including a range of consistently employed audio-visual aids. Discussions with a CI early in the training schedules followed each day session to review any departures from intended process- as confidence grew with the training these were replaced with accessibility of a CI to discuss any problems following the sessions.
The first intervention round for clinical services was developed as a two-day session, with the community services training planned as a separate two-day session in the same week. In addition to two consumer trainers employed by the project, trainers were sourced from clinical services for the clinical sessions and the community sector for the community sessions. This was anticipated as enabling the inclusion of specialist skills and experience in the delivery of training.
Training in the second round was subject to further modifications based on analyses of evaluations of the first round of training by both participants and trainers. The delivery of the intervention was modified to account for previously unknown restrictions on the ability of services to release staff for two days of training. Based on feedback from services, it was identified that attending two days of training for some teams was difficult. This was either due to the workload of the teams (specifically CAT teams) or the recent undertaking of organisational wide recovery training. In response to this the training was re-designed so that all material is covered in the first day of training, with more in-depth exploration and practice of the knowledge and skill on day two.

Feedback from the first round of training both through the structured feedback following training and from qualitative work led us to make several other modifications:
Training was restructured to allow half of the two days of training to be combined between the MHCSS sector and MH Staff. Feedback highlighted how the consumer role in leading training could be experienced as very challenging for some participants particularly if the consumer was experienced as critical of staff. Of course, being open to hearing criticisms from consumers about mental health care is a critical part of any transition to recovery-oriented practice so the training team worked very hard at considering this feedback in subsequent rounds. A key learning was that the introduction of the REACH coaching process needed to be deeply experiential. In particular, the training team formed the view that a critical element of the delivery was that the co-trainers as consumers and clinicians of other workers needed to embody the coaching principles in a fully authentic way. In alignment with a PDSA approach we took this on board as much as we could and adjusted the interactive style of the trainers for the second round. Specific focuses coaching based material was added with involvement of an additional trainer providing this particular perspective. Additions to the PULSAR Manual included sections providing information on Advance statements, Nominated persons and Risk and Recovery with additional references. Additional material was provided in Appendix 2 and the title changed from "Additional resources for understanding values" to "Additional resources for consumers' experiences". Additional web resources were added to Appendix 7 and the title changed from "Example of a relapse symptoms checklist" to "Care plans, and example of a relapse symptom checklist and other resources".

Appendix 3

Adverse events

At the commencement of the trial the forms of possible adverse events we anticipated included: 1) risk of distress by a participant during an interview; 2) issues related to disclosure of potential harm to self or others 3) risk of harm to staff. We developed an ethics protocol outlining the prevention and management of these risks which was approved by the governing HREC and our participant information and consent form for the face-to-face interviews outlined the potential risk of distress and what to do should it occur. Participants who were invited to complete the survey or undertake and interview were provided with written contact details for complaints, which was the manager of the governing HREC.
During the course of the project there were four complaints reported to the HREC. Three complaints related to QPR mailout (privacy concerns; receiving a letter but not a client of participating services; receiving a letter to a consumer who had died) which led to changes in procedures under direction of HREC as appropriate. One complaint related to the management of interview distress which led to updates to the staff training protocol and counselling provided to the staff member concerned.
No complaints were received that related to the REFOCUS-PULSAR intervention.
In addition to these complaints reported to HREC, during the trial one participant expressed suicidal ideation in a note attached to the return of a consent form. This was followed up by staff as per our ethics protocol to ensure their safety. The suicidal ideation was related to chronic psychiatric and medical symptoms and family conflict, not to participation in the project. In line with CATT advice, this issue was ultimately passed onto the police who took the participant to hospital, as the participant could not guarantee her safety.
Beyond the complaints process, we did not systematically collect any other adverse event information from consumers (such as deaths, hospitalisations etc.) because the intervention was a training intervention for staff rather than a clinical intervention for consumers so attribution of adverse events from clinical care in the context of the study, of which care guided by PULSAR-REFOCUS principles was but a part would not have been clear.

Appendix 4

Figure: Detailed CONSORT chart for stream-one

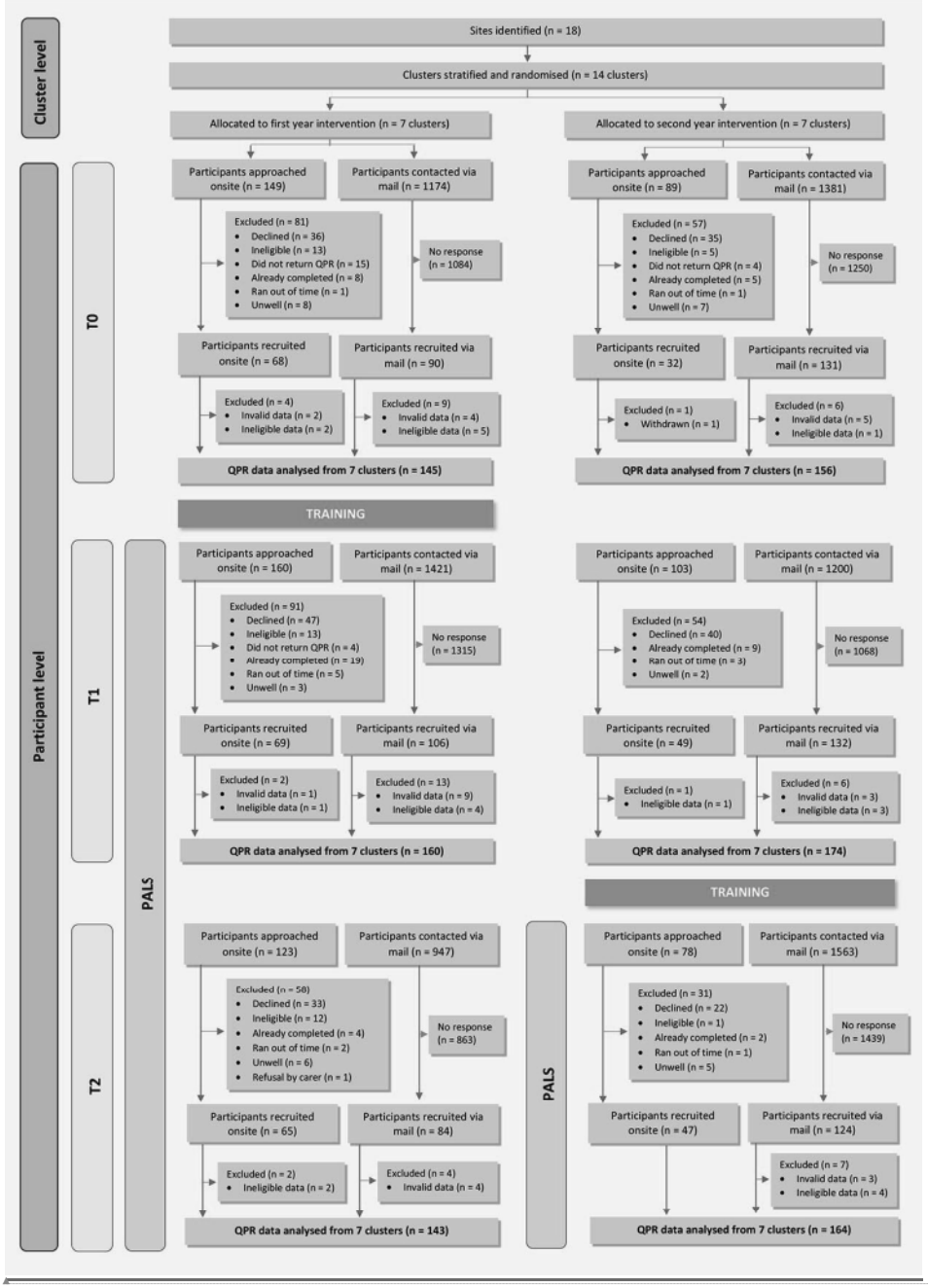


Figure S1: Detailed CONSORT chart stream-one

Notes. PALS = PULSAR Active Learning Sessions. Invalid data refers to data-based issues in the form of missing data or invalid responses. Ineligible data refers to participant-based issues – that is, the person providing the data did not meet the eligibility criteria for the study.

Formatted: Font: (Default) Calibri Light, 16 pt, Bold

Appendix 5

Details on profile of participants

Table 5.1 Stream-one trial QPR numbers (%) by Intervention Status, gender, age group, step, intervention status and demographics.

There were no significant differences between the control and intervention groups.

	Control	Intervention	Total
Distribution in specialist care			
N	475	467	942
(%)	(50·4·0)	(49·6)	(100)
Distribution in specialist care by Gender[1]			
Female	268 (56·4)	276 (59·1)	544 (57·7)
Male	203 (42·7)	187 (40·0)	390 (41·4)
Not listed	4 (0·8)	4 (0·9)	8 (0·8)
Distribution in specialist care by Age group			
17-30 years	104 (22·2)	125 (26·1)	229 (24·3)
30-49 years	243 (51·8)	229 (49·6)	472 (50·1)
50 years and over	122 (26·0)	108 (23·4)	230 (24·4)
Distribution in specialist care by Step Group intervention			
Step Group 1	145 (30·5)	300 (64·2)	445 (47·2)
Step Group 2	330 (69·5)	167 (35·8)	497 (52·8)
Distribution in specialist care by Intervention status (Ix)			
No Ix	475 (100)	0 (0·0)	475 (50·4)
Yes Ix	0 (0)	467 (100)	467 (49·6)
Distribution in specialist care by Country of birth			
Australia	345 (72·6)	345 (73·9)	690 (73·2)
Other[2]	125 (26·3)	118 (25·3)	243 (25·8)
Not listed	5 (1·1)	4 (0·9)	9 (1·0)
Distribution in specialist care by Year of arrival			
After 2000	27 (45·8)	32 (54·2)	59 (6·3)
Between 1981-2000	62 (58·5)	44 (41·5)	106 (11·3)
Before 1980	27 (54·0)	25 (46·0)	52 (5·2)
Not listed	12 (46·2)	14 (53·8)	26 (2·8)
Distribution in specialist care by Main language			
English	422 (88·8)	398 (85·2)	820 (87·0)
Other	32 (6·7)	40 (8·6)	72 (7·6)
Both English and Other	14 (2·9)	18 (3·9)	32 (3·4)
Not listed	7 (1·5)	11 (2·4)	18 (1·9)
Distribution in specialist care by Ethnicity (self-identified)			
Australian Non-Indigenous	223 (46·9)	237 (64·6)	460 (48·8)
Australian Indigenous	35 (7·4)	45 (12·3)	80 (8·5)
Other	176 (37·1)	167 (45·5)	343 (36·4)
Not listed	41 (8·6)	18 (4·9)	59 (6·3)
Distribution in specialist care by Duration of mental health service use			
Mean number of years	4·3	4·0	4·2
Median number of years	1·0	1·0	1·0
Range (years)	0-35	0-35	0-35
No. of people with <1 year at site	190 (40·0)	199 (42·6)	389 (42·3)
Mean number of months for those with <1 year at site	3·1	3·3	3·2
Median number of months for those with <1 year at site	3	3	3

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other_____

[2]Included 60 additional ethnic groups

Table 5.2 Stream-two trial QPR numbers (%) by timepoint, gender, age group, step, intervention status and demographics

	Timepoint			
	T0	T1	T2	Total
Distribution in specialist care by Timepoint				
N	71	129	73	273
(%)	(26·0)	(47·3)	(26·7)	(100)
Distribution in specialist care by Timepoint and Gender[1]				
Female	46 (64·8)	63 (48·8)	42 (57·5)	151 (55·3)
Male	25 (35·2)	64 (49·6)	31 (42·5)	120 (44)
Other	0 (0·0)	2 (1·6)	0 (0·0)	2 (0·7)
Distribution in specialist care by Timepoint and Age group				
17-30 years	26 (36·6)	23 (17·8)	13 (17·8)	62 (22·7)
30-49 years	30 (42·3)	67 (51·9)	35 (48)	132 (48·4)
50 years and over	15 (21·1)	38 (29·5)	24 (32·9)	77 (28·2)

Not listed	0 (0·0)	1 (0·8)	1 (1·4)	2 (0·7)
Distribution in specialist care by Timepoint and Step Group intervention				
Step Group 1	71 (100)	60 (46·5)	0 (0·0)	131 (48)
Step Group 2	0 (0·0)	69 (53·5)	73 (100)	142 (52)
Distribution in specialist care by Timepoint and Intervention status (Ix)				
No Ix	71 (100)	69 (53·5)	0 (0·0)	140 (51·3)
Yes Ix	0 (0·0)	60 (46·5)	73 (100)	133 (48·7)
Distribution in specialist care by Country of birth				
Australia	49 (69·0)	94 (72·9)	53 (72·6)	196 (71·8)
Other	22 (31·0)	35 (27·1)	20 (27·4)	77 (28·2)
Year of arrival in Australia				
After 2000	2 (9·1)	11 (18·3)	2 (10·0)	15 (19·5)
Between 1981-2000	12 (54·5)	14 (23·3)	10 (50·0)	36 (46·8)
Before 1980	6 (27·3)	10 (16·7)	7 (35·0)	23 (29·9)
Not listed	2 (9·1)	25 (41·7)	1 (5·0)	28 (36·4)
Distribution in specialist care by Main language spoken at home				
English	60 (84·5)	115 (89·2)	66 (90·4)	241 (88·3)
Other	11 (15·5)	14 (10·9)	7 (9·6)	32 (11·7)
Distribution in specialist care by Ethnicity (self-identified)				
Australian Non-Indigenous	38 (53·5)	74 (57·4)	43 (58·9)	155 (56·8)
Australian Indigenous	2 (2·8)	2 (1·6)	2 (2·7)	6 (2·2)
Other	20 (28·2)	50 (38·8)	19 (26)	89 (32·6)
Not listed	11 (15·5)	3 (2·3)	9 (12·3)	23 (8·4)
Other category (multiple responses could be listed)				
British (English, Irish, Walsh, Scottish)	1 (5·0)	17 (34·0)	5 (26·3)	23 (25·8)
European (Italian, Greek, Bosnian, Dutch, German)	7 (35·0)	20 (40·0)	5 (26·3)	32 (36·0)
New Zealander/Maori	2 (10·0)	3 (6·0)	4 (21·1)	9 (10·1)
Middle Eastern (Afghan)	2 (10·0)	0 (0·0)	0 (0·0)	2 (2·2)
South East Asian (Burmese, Chinese, Indian, Cambodian, Sri Lankan, Vietnamese)	8 (16·0)	5 (10·0)	5 (26·3)	18 (20·2)
Other (participant selected "other")	0 (0·0)	5 (10·0)	0 (0·0)	5 (5·6)
Distribution in specialist care by Duration of mental health service use				
Mean number of years	11·0	13·2	13·1	12·6
Median number of years	9·0	11·0	11·0	10·0
Range (years)	1-40	1-33	1-40	1-40
No. of people with <1 year at site	0	0	0	0
Duration of current service use				
Mean number of years	4·6	5·8	7·2	5·8
Median number of years	3·0	3·0	4·5	3·0
Range (years)	0-23	1-22	1-32	0-32
No. of people with <1 year at site	1	0	0	1
Median number of months for those with <1 year at site	3	-	-	3
Distribution in specialist care by Marital status				
Single	48 (67·6)	67 (51·9)	33 (45·2)	148 (54·2)
Married	6 (8·5)	21 (16·3)	16 (21·9)	43 (15·8)
DeFacto	3 (4·2)	8 (6·2)	1 (1·4)	12 (4·4)
Separated	4 (5·6)	15 (11·6)	10 (13·7)	29 (10·6)
Divorced	10 (14·1)	14 (10·9)	10 (13·7)	34 (46·6)
Widowed	0 (0·0)	1 (0·8)	2 (2·7)	3 (4·1)
Other	0 (0·0)	3 (2·3)	1 (1·4)	4 (5·5)
Distribution in specialist care by Child status				
Yes	33 (46·5)	64 (49·6)	41 (56·2)	138 (50·5)
No	38 (53·5)	65 (50·4)	32 (43·8)	135 (49·5)
Number of children living at home				
0	16 (48·5)	32 (50·0)	16 (39·0)	64 (46·4)
1	11 (33·3)	17 (26·6)	12 (29·3)	40 (29·0)
2	4 (12·2)	10 (15·6)	9 (22·0)	23 (16·7)
3	0 (0·0)	1 (1·6)	3 (7·3)	4 (2·9)
4	1 (3·0)	1 (1·6)	0 (0·0)	2 (1·4)
5-6	1 (3·0)	1 (1·6)	0 (0·0)	2 (1·4)
Not listed	0 (0·0)	2 (3·1)	1 (2·4)	3 (2·2)
Distribution in specialist care by Living situation (multiple responses could be selected)				
Living with parents	16 (22·5)	16 (12·4)	15 (20·5)	47 (17·2)
Living with siblings	4 (5·6)	8 (6·2)	4 (5·5)	16 (5·9)
Living with a partner	7 (9·9)	31 (24·0)	15 (20·5)	53 (19·4)
Living with children	15 (21·1)	26 (20·2)	19 (26·0)	60 (22·0)
Living with friends	4 (5·6)	7 (5·4)	2 (2·7)	13 (4·8)
Living in shared accommodation	5 (7·0)	14 (10·9)	5 (6·8)	24 (8·8)
Living in crisis accommodation	3 (4·2)	3 (2·3)	1 (1·4)	7 (2·6)
Living in support housing	8 (11·3)	11 (8·5)	9 (12·3)	28 (10·3)
Living alone	23 (32·4)	30 (23·3)	20 (27·4)	73 (26·7)
Homeless	3 (4·2)	3 (2·3)	0 (0·0)	6 (2·2)
Other	4 (5·6)	10 (7·8)	1 (1·4)	15 (5·5)
Distribution in specialist care by Education level				
None	0 (0·0)	0 (0·0)	1 (1·4)	1 (0·4)
Primary school	2 (2·8)	4 (3·1)	0 (0·0)	6 (2·2)

Secondary school (≤ yr 10)	25 (35·2)	39 (27·9)	21 (28·8)	85 (31·1)
Secondary school (yr 11)	12 (16·9)	30 (23·3)	14 (19·2)	56 (20·5)
Secondary school (yr 12)	27 (38·0)	56 (43·4)	37 (50·7)	120 (44·0)
Not listed	5 (7·0)	0 (0·0)	0 (0·0)	5 (1·8)
Distribution in specialist care by Highest qualification				
Certificate I	2 (2·8)	1 (0·8)	1 (1·4)	4 (1·5)
Certificate II	3 (4·2)	6 (4·7)	4 (5·5)	13 (4·8)
Certificate III	15 (21·1)	19 (14·7)	13 (17·8)	47 (17·2)
Certificate IV	7 (9·9)	15 (11·6)	10 (13·7)	32 (11·7)
Diploma	9 (12·7)	11 (8·5)	10 (13·7)	30 (11·0)
Advanced Diploma	1 (1·4)	1 (0·8)	1 (1·4)	3 (1·1)
Associate Degree	0 (0·0)	1 (0·8)	0 (0·0)	1 (0·4)
Bachelor Degree	1 (1·4)	13 (10·1)	7 (9·6)	21 (7·7)
Bachelor Honours Degree	0 (0·0)	3 (2·3)	1 (1·4)	4 (1·5)
Graduate Diploma	0 (0·0)	2 (1·6)	2 (2·7)	4 (1·5)
Masters (research)	0 (0·0)	1 (0·8)	0 (0·0)	1 (0·4)
Masters (coursework)	1 (1·4)	3 (2·3)	1 (1·4)	5 (1·8)
Doctoral	1 (1·4)	1 (0·8)	0 (0·0)	2 (0·7)
Other	3 (2·8)	11 (8·5)	4 (5·5)	18 (6·6)
Not listed	28 (39·4)	41 (31·8)	19 (26·0)	85 (31·1)

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other.

Table 5.3 Stream-two trial QPR numbers (%) by Intervention Status, gender, age group, step, intervention status and demographics.

There were no significant differences between the control and intervention groups.

	Control	Intervention	Total
Distribution in specialist care			
N	140	133	273
(%)	(51·3)	(48·7)	(100)
Distribution in specialist care by Gender[1]			
Female	80 (57·1)	71 (53·4)	151 (55·3)
Male	59 (42·1)	61 (45·9)	120 (44)
Not listed	1 (0·7)	1 (0·8)	2 (0·7)
Distribution in specialist care by Age group			
17-30 years	35 (25·0)	27 (20·3)	62 (22·7)
30-49 years	65 (46·4)	67 (50·4)	132 (48·4)
50 years and over	39 (27·9)	38 (28·6)	77 (28·2)
Not listed	1 (0·7)	1 (0·8)	2 (0·7)
Distribution in specialist care by Step Group intervention			
Step Group 1	71 (50·7)	60 (45·1)	131 (48·0)
Step Group 2	69 (49·3)	73 (54·9)	142 (52·0)
Distribution in specialist care by Intervention status (Ix)			
No Ix	140 (100)	0 (0·0)	140 (50·4)
Yes Ix	0 (0·0)	133 (100)	133 (49·6)
Distribution in specialist care by Country of birth			
Australia	98 (70·0)	98 (73·7)	196 (71·8)
Other[2]	42 (30·0)	35 (26·3)	77 (28·2)
Year of arrival in Australia			
After 2000	8 (19·0)	7 (20·0)	15 (19·5)
Between 1981-2000	20 (47·6)	16 (45·7)	36 (46·8)
Before 1980	12 (28·6)	11 (31·4)	23 (29·9)
Not listed	2 (4·8)	1 (2·9)	3 (3·9)
Distribution in specialist care by Main language			
English	124 (88·6)	117 (88·0)	241 (88·3)
Other	16 (11·4)	16 (12·0)	32 (11·7)
Distribution in specialist care by Ethnicity (self-identified)			
Australian Indigenous	2 (1·4)	4 (3·0)	6 (2·2)
Australian Non-Indigenous	77 (55·0)	78 (58·6)	155 (56·8)
Other	54 (38·6)	35 (26·3)	89 (32·6)
Not listed	7 (5·0)	16 (12·0)	23 (8·4)
Other category (multiple responses could be listed)			
British (English, Irish, Walsh, Scottish)	13 (24·1)	10 (28·6)	23 (25·8)
European (Italian, Greek, Bosnian, Dutch, German)	22 (40·7)	10 (28·6)	32 (36·0)
New Zealander/Maori	4 (7·4)	5 (14·3)	9 (10·1)
Middle Eastern (Afghan)	2 (3·7)	0 (0·0)	2 (2·2)
South East Asian (Burmese, Chinese, Indian, Cambodian, Sri Lankan, Vietnamese)	13 (24·1)	10 (28·6)	23 (25·8)
Other (participant selected "other")	0 (0·0)	0 (0·0)	0 (0·0)
Distribution in specialist care by Duration of mental health service use			
Mean number of years	13·1	12·0	12·6
Median number of years	10·0	10·0	10·0
Range (years)	1-40	1-40	1-40

No. of people with <1 year at site10 (28·6)	0	0	0
Duration of current service use			
Mean number of years	5·2	6·5	5·8
Median number of years	3·0	4·0	3·0
Range (years)	0-23	1-32	0-32
No. of people with <1 year at site	1	0	1
Number of months for those with <1 year at site	3	-	3
Distribution in specialist care by Marital status			
Single	77 (55·0)	71 (53·4)	148 (54·2)
Married	20 (14·3)	23 (17·3)	43 (15·8)
DeFacto	10 (7·1)	2 (1·5)	12 (4·4)
Separated	15 (10·7)	14 (10·5)	29 (10·6)
Divorced	17 (12·1)	17 (12·8)	34 (12·5)
Widowed	0 (0·0)	3 (2·3)	3 (1·1)
Other	1 (0·7)	3 (2·3)	4 (1·5)
Distribution in specialist care by Child status			
Yes	76 (54·3)	62 (46·6)	138 (50·5)
No	64 (45·7)	71 (53·4)	135 (49·5)
Number of children living at home			
0	39 (51·3)	25 (40·3)	64 (46·4)
1	11 (14·5)	21 (33·9)	32 (23·2)
2	12 (15·8)	11 (17·7)	23 (16·7)
3	2 (2·6)	3 (4·8)	5 (3·6)
4	2 (2·6)	0 (0·0)	2 (1·4)
5-6	1 (1·3)	0 (0·0)	1 (0·7)
Not listed	37 (48·7)	2 (3·2)	39 (28·2)
Distribution in specialist care by Living situation (multiple responses could be selected)			
Living with parents	21 (15·0)	26 (19·5)	47 (17·2)
Living with siblings	8 (5·7)	8 (6·0)	16 (5·9)
Living with a partner	29 (20·7)	24 (18·0)	53 (19·4)
Living with children	31 (22·1)	29 (21·8)	60 (22·0)
Living with friends	6 (4·3)	7 (5·3)	13 (4·8)
Living in shared accommodation	12 (8·6)	12 (9·0)	24 (8·8)
Living in crisis accommodation	6 (4·3)	1 (0·8)	7 (2·6)
Living in support housing	12 (8·6)	16 (12·0)	28 (10·3)
Living alone	38 (27·1)	35 (12·8)	73 (26·7)
Homeless	4 (2·9)	2 (1·5)	6 (2·2)
Other	12 (8·6)	3 (2·3)	15 (5·5)
Distribution in specialist care by Education level			
None	0 (0·0)	1 (0·8)	1 (0·4)
Primary school	4 (2·9)	2 (1·5)	6 (2·2)
Secondary school (≤ yr 10)	46 (32·9)	39 (29·3)	85 (31·1)
Secondary school (yr 11)	30 (21·4)	26 (19·5)	56 (42·1)
Secondary school (yr 12)	55 (39·3)	65 (48·9)	120 (44·0)
Not listed	5 (3·6)	0 (0·0)	5 (1·8)
Distribution in specialist care by Highest qualification			
Certificate I	2 (1·4)	2 (1·5)	4 (1·5)
Certificate II	7 (5·0)	6 (4·5)	13 (4·8)
Certificate III	26 (18·6)	21 (15·8)	47 (17·2)
Certificate IV	14 (10·0)	18 (13·5)	32 (11·7)
Diploma	18 (12·9)	12 (9·0)	30 (11·0)
Advanced Diploma	1 (0·7)	2 (1·5)	3 (1·1)
Associate Degree	0 (0·0)	1 (0·8)	1 (0·4)
Bachelor Degree	8 (5·7)	13 (9·8)	21 (7·7)
Bachelor Honours Degree	1 (0·7)	3 (2·3)	4 (1·5)
Graduate Diploma	2 (1·4)	2 (1·5)	4 (1·5)
Masters (research)	1 (0·7)	0 (0·0)	1 (0·4)
Masters (coursework)	2 (1·4)	3 (2·3)	5 (1·8)
Doctoral	1 (0·7)	1 (0·8)	2 (0·7)
Other	9 (6·4)	9 (6·8)	18 (6·6)
Not listed	48 (34·3)	40 (30·1)	88 (32·2)

Note. Where cell sizes are less than 5 at any timepoint for a given characteristic, data were pooled to ensure confidentiality.

[1]Gender was determined by asking participants: "Which gender do you identify with?" with options being Male, Female, Other.

Appendix 6

Stream Two Blindness findings

Process and findings

At the end of each PULSAR stream-two interview, the interviewer recorded if they thought the participant had come from a site that had received the intervention in year one (step-one) or year two (step-two). Interviewers were required to make a guess if they had no thought about the site intervention status. The null hypothesis here is that the observed proportion of correct guesses is 0·5 (i.e. half-and-half). Binomial probability theory tells us the probability of a type 1 error (the incorrect rejection of a true null hypothesis), which is shown in Table 1.

Table 6 Accuracy of interviewer guesses regarding study site in stream-two

	Number of interviews	Correct guess (number)	Correct guess (%)	Probability of Type 1 error: p value
T0	74	49	66·2%	0·004
T1	45	17	37·8%	0·964
T2	74	41	55·4%	0·208

At the first timepoint T0, the proportion of interviewers correctly guessing the intervention allocation at sites (66·2%) was significant, with p-value<0·01. On review of these results during the study the team considered they may have been influenced by the non-random order of selection options, with interviewers possibly being more likely to tick the first box (which was consistently the correct option). So for the later timepoints T1 and T2, the options for interviewers to select were randomised. At both of these timepoints, the proportion of interviewers correctly guessing the intervention allocation at sites was no different to chance (T1 was 37·8%, p-value>0·9; T2 was 55·4%, p-value>0·2).

Summary and conclusions

Assessment of blindness for stream-two interviews indicated that at T0, the proportion of interviewers correctly guessing site intervention allocation (66·2%) was significant, p<0·01. This result was possibly influenced by non-random ordering of selection options and as options were randomised for T1 and T2, the proportion interviewers correct guessing of site intervention allocation was no different from chance (T1: 37·8%, p>0·9; T2: 55·4%, p>0·2). We conclude it is unlikely through the course of the project that interview bias would represent a significant bias to findings.

Appendix 7

Main model building for stream-one QPR outcome

The model building process is shown in models 1.1 to 1.4, where model 1.4 is the final main model referred to in the manuscript. All models below have the cluster variable specified as random. Model building begins with fixed factors of timepoint and intervention status in Model 1.1. Then in model 1.2, fixed effects of sex and age group are added. Then in model 1.3, added is the fixed effect variable for sector (PMHS; MHCSS). Finally, model 1.5 has same variables as model 1.7 plus 'step' group as fixed variable.

Model 1.1 Stream-one QPR mixed model with fixed factors of timepoint and intervention status, and clusters as random.

Number of obs=942.

*** p<0·01, ** p<0·05, * p<0·1

		b	se	z	p value	ll	ul
Timepoint							
	T1	-3·59	1·13	-3·17	·002***	-5·81	-1·37
	T2	-4·78	1·60	-2·99	·003***	-7·92	-1·65
Intervention status							
	yes	4·15	1·54	2·69	·007***	1·13	7·18

Model Adjusted QPR means

		Model adj. QPR mean	Std.Err.	[95%Conf.	Interval]
Timepoint					
	T0	57·14	1·25	54·70	59·58
	T1	53·55	1·34	50·92	56·18
	T2	52·35	1·39	49·63	55·07
Intervention status*					
	0	52·25	1·46	49·37	55·12
	1	56·40	1·25	53·95	58·84

*The mean difference between treatment and control groups was 4·2 (95% Confidence interval: 1·1 – 7·2).

Model 1.2. Stream-one QPR mixed model with fixed factors of sex, age-group, time and intervention status, and

clusters as random.
Number of obs= 942.
*** p<0·01, ** p<0·05, * p<0·1.

		b	se	z	p value	ll	ul
Sex							
	Female	-·86	1·05	-0·82	0·414	-2·92	1·20
Age group							
	2	-·88	·86	-1·02	0·308	-2·56	0·81
	3	-3·40	·93	-3·67	0·001***	-5·21	-1·58
Timepoint							
	T1	-3·20	1·07	-2·99	0·003***	-5·29	-1·11
	T2	-4·19	1·56	-2·69	0·007***	-7·25	-1·14
Intervention status							
	Yes	3·74	1·56	2·36	0·018**	·63	6·85

Model adjusted QPR means

		QPR mean	Std.Err.	[95%Conf.	Interval]
Sex					
	Male	54·82	1·29	52·29	57·35
	Female	53·96	1·25	51·51	56·41
Age group					
	1	55·60	1·12	53·40	57·80
	2	54·72	1·22	52·34	57·10
	3	52·20	1·50	49·26	55·15
Timepoint					
	0	56·82	1·18	54·52	59·13
	1	53·63	1·38	50·93	56·33
	2	52·63	1·47	49·74	55·52
Intervention status*					
	0	52·47	1·53	49·46	55·47
	1	56·20	1·25	53·75	58·66

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0.7 – 6·8).

Model 1.3. Stream-one QPR mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.
Number of obs=942.
*** p<0·01, ** p<0·05, * p<0·1.

		b	se	z	p value	ll	ul
Sex							
	Female	-0·818	1·05	-·78	·434	-2·87	1·23
Age group							
	2	-0·93	0·85	-1·09	·274	-2·59	0·74
	3	-3·43	0·91	-3·77	0·001***	-5·21	-1·65
Timepoint							
	T1	-3·20	1·07	-2·99	·003***	-5·30	-1·10
	T2	-4·18	1·56	-2·67	·008***	-7·24	-1·11
Intervention status							
	Yes	3·72	1·64	2·27	·023**	0·51	6·92
Sector							
	2	-1·71	2·12	-0·81	·418	-5·87	2·44

Model adjusted QPR means

		QPR mean	Std.Err.	[95%Conf.	Interval]
Sex					
	1	54·85	1·35	52·20	57·49
	2	54·03	1·22	51·63	56·42
Age group					
	1	55·68	1·13	53·46	57·90
	2	54·75	1·24	52·31	57·19
	3	52·25	1·48	49·35	55·15
Timepoint					
	0	56·87	1·203	54·51	59·23
	1	53·67	1·42	50·885	56·45
	2	52·69	1·44	49·87	55·51
Intervention status					
	0	52·53	1·57	49·45	55·60
	1	56·24	1·26	53·77	58·72
Sector					
	1	55·04	1·74	51·63	58·45
	2	53·39	1·23	50·91	55·75

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0·5 – 6·9).

Model 1.4. Stream-one QPR mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.

Number of obs=942.

*** p<0·01, ** p<0·05, * p<0·1.

		b	se	z	p value	ll	ul
Sex							
	Female	-0·81	1·04	-0·79	·431	-2·84	1·21
Age group							
	2	-0·94	0·88	-1·07	·285	-2·65	0·78
	3	-3·4	0·91	-3·78	0·001***	-5·22	-1·66
Timepoint							
	T1	-3·22	1·02	-3·16	·002***	-5·22	-1·22
	T2	-4·22	1·50	-2·82	·005***	-7·15	-1·29
Intervention status							
	Yes	3·76	1·31	2·87	·004***	1·20	6·33
Sector							
	2	-1·72	2·12	-0·81	·418	-5·87	2·43
Step group							
	2	0·15	2·08	0·07	·943	-3·93	4·22

Model adjusted QPR means

		QPR mean	Std.Err.	[95%Conf.	Interval]
Sex					
	1	54·85	1·36	52·18	57·51
	2	54·03	1·23	51·63	56·44
Age Group					
	1	55·69	1·17	53·39	57·99
	2	54·75	1·25	52·31	57·20
	3	52·25	1·48	49·34	55·16
Timepoint					
	0	56·89	1·25	54·45	59·34
	1	53·67	1·43	50·86	56·48
	2	52·67	1·38	49·97	55·37
Intervention status					
	0	52·51	1·46	49·65	55·37
	1	56·27	1·23	53·87	58·67
Sector					
	1	55·05	1·75	51·62	58·47
	2	53·33	1·24	50·91	55·75
Step					
	1	54·29	1·42	51·52	57·07
	2	54·44	1·69	51·12	57·76

*The mean difference between treatment and control groups was 3·7 (95% Confidence interval: 0·5 – 6·8).

Appendix 8

Interaction term model for stream-one QPR outcome

The model building process is shown in models 1.5 to 1.8, where model 1.8 is the final interaction term model referred to in the manuscript. All models below have the cluster variable specified as random. Reference groups in models can be arbitrary, and were selected based on the lowest QPR means at the specified timepoint. In Appendix 12 are same models but with the first group as the reference. Model building begins with the interaction item of time and intervention status in Model 1.5. Then in model 1.6, fixed effects of sex and age group are added. Then in model 1.7, added is the fixed effect variable for sector (PMHS; MHCSS). Finally, model 1.8 has same variables as model 1.7 plus 'step' group as fixed variable.

Model 1.5. Stream-one QPR has interaction item of time and intervention status.

Timepoint	Intervention	Coefficient	Robust Std Err.	P>|z|	95% CI	
T0	No	3·59	1·13	0·002	1·37	5·81
T1	No	Reference				
T1	Yes	4·15	1·54	0·007	1·13	7·18
T2	Yes	2·96	1·39	0·030	0·24	5·68

Number of obs=942.

Model adjusted QPR means

			Model adjusted statistics	
Timepoint	Intervention	QPR raw data	QPR mean	95% CI

			mean			
T0		No	54·7	55·1	52·7	57·5
T1		No	51·5	51·5*	48·2	54·8
	Yes	Note. 55·3Wald test for interaction terms: $\chi^2(3) = 11·05$. p=0·011, providing evidence of significant interaction between the time and intervention variables.	55·6*	52·9	58·4	
T2		Yes	53·9	54·4	52·1	56·8

*The mean difference between treatment and control groups at year 1 was 4·2 (95% Confidence interval: 1·1 – 7·2).

(B) **Model 1.6. Stream-one QPR 2 has interaction itemfixed variables of time and, intervention status, and fixed variables of age-group (<30; 30-49; 50 years and over) and sex (Male/Female).)**

Number of obs=942

		Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex						
	Male			Reference		
	Female	-0·86	1·05	0·41	-2·92	1·20
Age Category						
	17-29			Reference		
	30-49	-0·88	0·86	0·31	-2·56	0·81
	50-75	-3·39	0·92	<0·001	-5·21	-1·58
TimepointTime point	**Intervention**					
T0	No	3·20	1·07	0·003	1·10	5·29
T1	No			Reference		
T1	Yes	3·74	1·59	0·02	0·63	6·85
T2	Yes	2·74	1·35	0·04	0·09	5·34

Model adjusted QPR means

				Model adjusted statistics	
Timepoint		**Intervention**	**QPR raw data mean**	**QPR mean**	**95% CI**
T0	No	54·7	55·0	52·6	Note. 57·3Wald test for interaction terms: $\chi^2(3) = 9·45$. p=0·024. The Wald test provided evidence of significant interaction between the time and intervention variables.
T1		No	51·5	51·8*	48·4 55·2
		Yes	55·3	55·5*	52·7 58·3
T2		Yes	53·9	54·5	52·1 56·9

* The mean difference between treatment and control groups at year 1 was 3·7 (95% Confidence interval: 0·6 –

6·8).

(C) **Model 1.7. Stream-one QPR Model3 has same variables as Model 1.62 plus sector (PMHS; MHCSS) as fixed.**

Number of obs=942

			Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex							
		Male (reference)					
		Female	-0·76	1·05	0·47	-2·81	1·29
Age Category							
		17-29 (reference)					
		30-49	-0·90	0·84	0·28	-2·55	0·75
		50-75	-3·37	0·91	0·00	-5·15	-1·59
TimepointTime point	**Intervention**	**Sector**					
T0	No	**PMHS**	3·56	1·47	0·02	0·68	6·43
T0	No	**MHCSS**	2·58	2·65	0·33	-2·61	7·78
T1	No	**PMHS**	(reference)				
T1	No	**MHCSS**	0·03	3·23	0·99	-6·30	6·35
T1	Yes	**PMHS**	3·99	2·23	0·07	-0·38	8·37
T1	Yes	**MHCSS**	3·33	2·64	0·31	-1·85	8·51
T2	Yes	**PMHS**	4·30	1·85	0·02	0·67	7·93
T2	Yes	**MHCSS**	0·32	2·89	0·91	-5·35	5·98

Note. Wald test for interaction terms: $\chi^2(7) = 19{\cdot}52$. p=0·007. The Wald test provided evidence of significant interaction between the time, intervention and sector variables.

~~(D)~~ **Model adjusted QPR means**

Sector	Timepoint	Intervention	QPR raw data mean	Model adjusted statistics: QPR mean	95% CI	
PMHS	T0	No	55·0	55·4	51·6	59·2
	T1	No	51·1	51·8	47·0	56·7
		Yes	55·4	55·8	51·4	60·2
	T2	Yes	55·1	56·1	53·0	59·2
MHCSS	T0	No	54·3	54·4	52·6	56·3
	T1	No	52·3	51·9	47·8	55·9
		Yes	54·8	55·2	53·3	57·1
	T2	Yes	52·1	52·2	49·2	55·1

Model 1.8. Stream-one QPR Model~~4~~ has same variables as model 1.7~~3~~ plus 'step' group as fixed variable. Model ~~4~~ also examined interactions between four variables (sector, step, time and intervention).

Number of obs=942

				Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex								
			Male (reference)					
			Female	-0·76	1·06	0·47	-2·85	1·32
Age Category								
			17-29 (reference)					
			30-49	-0·87	0·87	0·32	-2·56	0·83
			50-75	-3·37	0·94	0·00	-5·22	-1·52
~~Timepoint~~Time point	**Intervention**	**Step Group**	**Sector**					
T0	No	1	PMHSPMHS	5·83	3·40	0·09	-0·84	12·49
T0	No	2	PMHS	2·36	1·50	0·12	-0·58	5·30
T1	No	2	PMHS	(reference)				
T1	Yes	1	PMHS	5·02	4·21	0·23	-3·23	13·27
T2	Yes	1	PMHS	4·48	4·37	0·31	-4·09	13·04
T2	Yes	2	PMHS	4·92	1·65	0·00	1·68	8·16
T0	No	1	MHCSS	1·99	3·32	0·55	-4·52	8·50
T0	No	2	MHCSS	4·14	3·39	0·22	-2·50	10·78
T1	No	2	MHCSS	1·29	4·11	0·76	-6·77	9·34
T1	Yes	1	MHCSS	3·11	3·23	0·34	-3·22	9·43
T2	Yes	1	MHCSS	-0·29	3·16	0·93	-6·47	5·90
T2	Yes	2	MHCSS	1·94	4·20	0·65	-6·29	10·17

~~Note. Wald test for the interaction terms: $\chi^2(11) = 372{\cdot}52$. p<0·0001. The Wald test provided evidence of significant interactions.~~

Table 4 Model adjusted QPR means. All models had specified cluster as random effect see T3 comments

(A) Model 1 has fixed variables of time and intervention status.

					Model adjusted QPR means statistics			
Sector Time Point	Timepoint		Step group	Intervention	QPR raw data mean	QPR mean	95% CI	Pre/post intervention diff. *Significant
T0	No	54·7	55·1		52·7	57·5		
T1	No	51·5	51·5*		48·2	54·8		
	Yes	55·3	55·6*		52·9	58·4		
T2	Yes	53·9	54·4		52·1	56·8		

*The mean difference between treatment and control groups at year 1 was 4·2 (95% Confidence interval: 1·1 – 7·2).

(B) Model 2 has fixed variables of time, intervention status, age group and sex. Next, added fixed variables of age (<30; 30-49; 50 years and over) and sex (male/female) into model.

			Model adjusted statistics	
Time Point	Intervention	QPR raw data mean	QPR mean	95% CI
T0	No	54·7	55·0	52·6 – 57·3
T1	No	51·5	51·8*	48·4 – 55·2
	Yes	55·3	55·5*	52·7 – 58·3
T2	Yes	53·9	54·5	52·1 – 56·9

* The mean difference between treatment and control groups at year 1 was 3·7 (95% Confidence interval: 0·6 – 6·8).

(C) Model 3 has same variables as model 2 plus sector (PMHS; MHCSS) as fixed.

				Model adjusted statistics	
Sector	Time Point	Intervention	QPR raw data mean	QPR mean	95% CI
PMHS	T0	No	55·0	55·4	51·6 – 59·2
	T1	No	51·1	51·8	47·0 – 56·7
		Yes	55·4	55·8	51·4 – 60·2
	T2	Yes	55·1	56·1	53·0 – 59·2
MHCSS	T0	No	54·3	54·4	52·6 – 56·3
	T1	No	52·3	51·9	47·8 – 55·9
		Yes	54·8	55·2	53·3 – 57·1
	T2	Yes	52·1	52·2	49·2 – 55·1

(D) Model 4 has same variables as Model 3 plus 'step' group as fixed variable. Model 4 also examined interactions between four variables (sector, step, time and intervention).

					Model adjusted statistics			
Sector	Time Point	Step group	Intervention	QPR raw data mean	QPR mean		95% CI	Pre/post intervention diff. *Significant
PMHS	T0	1	*No*	*57·1*	*57·2*	*54·6*	*59·7*	
	T1	1	Yes	55·4	56·4	50·9	61·8	-0·8 (z-score=0·5, p=0·64)
	T2	1	Yes	54·8	55·8	49·9	61·7	
PMHS	T0	2	*No*	*53·0*	*53·7*	*47·3*	*60·1*	
	T1	2	*No*	*51*·1	*51·3*	*45·2*	*57·5*	
	T2	2	Yes	55·4	56·2	53·0	59·5	4·9 (z-score=3·0, p=0·003)*
MHCSS	T0	1	*No*	*53·3*	*53·3*	*51·0*	*55·6*	
	T1	1	Yes	54·8	54·4	52·7	56·2	1·1 (z-score=2·7, p=0·006)*
	T2	1	Yes	51·3	51·0	50·0	52·0	
MHCSS	T0	2	*No*	*55·1*	*55·5*	*53·0*	*58·0*	
	T1	2	*No*	*52*·3	*52·6*	*47·4*	*57·8*	
	T2	2	Yes	52·7	53·2	47·8	58·7	0·7 (z-score=1·22, p=0·22)

Appendix 9

Perceived need for care findings

Instrumentation

The Perceived Need for Care Questionnaire is an interviewer administered questionnaire that in the form here used classifies seven forms of need:

1. Information about mental illness, its treatments and available services. (Information)
2. Medicine or tablets. (Medicines)
3. Counselling or talking therapy. (Counselling)
4. Practical issues such as housing or money issues. (Practical)
5. Help to improve the ability to work or use time in other ways. (Time use)
6. Help to improve the ability to look after themselves in their home. (Self-care)
7. Help to meet people for support and company (Company)

Through a branching conversationally styled question structure these needs are identified as judged by the participant to fall into four perceived need categories: no need, unmet need, partially met need, or met need.

Hypotheses

Here we examine three hypotheses, H1-H3: H 1: People in intervention as an outcome of more comprehensive assessment would identify more needs: H 2: People in intervention would be more likely to identify needs where present as met and less likely to identify them as unmet. H 3: H 2 would apply especially in more personal recovery than clinical goals areas, so here items 4-7.

Results

Table 9 Need Categories assessed with the Perceived Need for Care Questionnaire as associated with intrervention status

Perceived Need category		PULSAR-REFOCUS Intervention status	No need (a)	Unmet need (b)	Partially met need (c)	Met need (d)	Proportion of all needs met (d/(b+c+d)	Proportion of all needs unmet b/(b+c+d)
1	Information	Control	22	14	35	66	57·4%	12%
		Intervention	10	23	26	66	57·4%	20%
2	Medicines	Control	6	0	26	105	80·2%	0%
		Intervention	2	4	17	100	82·6%	3%
3	Counselling	Control	11	13	38	75	59·5%	10%
		Intervention	11	12	41	61	53·5%	11%
4 -	Practical	Control	50	35	18	34	39·1%	40%
		Intervention	42	32	15	41	46·6%	36%
5 -	Time use	Control	48	38	14	35	40·2%	44%
		Intervention	41	31	13	42	48·8%	36%
6 -	Self-care	Control	56	32	13	38	45·8%	39%
		Intervention	42	29	13	42	50·0%	35%
7 -	Company -	Control	38	34	16	46	47·9%	35%
		Intervention	37	26	23	39	44·3%	30%

Here, given the categorical nature of the data, smaller sample sizes than for primary outcome variables, and without expectation of this part of the study being fully powered, we have kept statistical analyses very simple.

H 1 - People in intervention as an outcome of better assessment would identify more needs: Here we find people in the in intervention group identified a perceived need in 696 of 881 invitations to do so 79% while among control participants this proportion was 725/956 or 76%. A two sample test of proportions result gives a z-statistic = -1·54, p=0·0622, so in the marginal significance range of 0·05-0·10.

H 2 - People in intervention would be more likely to identify needs where present as met and less likely to identify these as unmet. Here, comparisons favour the intervention 8:5 with one tie. In 13 items, 8 favouring the intervention will occur by chance with a probability of 0·157 i.e. p=0·157 so here the probability of type I error in relation to the proposition that more needs will be identified in intervention group participants is 0·157 (here p>0·10 NS).

H 3: H 2 would apply especially in more personal recovery than clinical goals areas, so here items 4-7. Here comparisons favour the intervention 7:1. In 8 items, 7 favouring the intervention will occur by chance with a probability of 0·031. So here the p-value is 0·031 i.e. probability of type I error is 0·031 (here p< 0·05).

Conclusion

While noting the limitations of the analyses, two of the three hypotheses receive some support, one with p<0·05 and another with 0·05 < p < 0·1 while the third is favoured in terms of direction of findings, though not significantly so. Considered in the context of the overall set of measures we would rate the PNCQ findings as overall favourable for the intervention condition over controls.

Appendix 10

Stream-two models of the QPR

Model 2.1 Stream-two QPR mixed model with fixed factors of time and intervention status, and clusters as random.

Number of obs=269.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-0·50	2·79	-0·18	0·86	-5·95	4·96
T2	-2·74	3·94	-0·70	0·49	-10·45	4·98
Intervention status						
1	2·446	2·90	0·84	0·40	-3·23	8·12

Model Adjusted QPR means

	QPR mean	Std.Err.	[95%Conf.	Interval]
Timepoint				
T0	54·50	1·44	51·68	57·31
T1	54·00	1·47	51·12	56·88
T2	51·76	2·77	46·33	57·18
Intervention status				
0	52·34	2·29	47·84	56·83
1	54·78	0·93	52·96	56·60

Model 2.2 Stream-two QPR mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs=265.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·66	1·82	-1·47	0·143	-6·22	0·90
Age group						
30-49 years	0·98	1·79	0·55	0·585	-2·53	4·49
50 years and over	-3·01	2·11	-1·43	0·153	-7·14	1·12
Timepoint						
T1	-0·88	2·75	-0·32	0·749	-6·26	4·51
T2	-2·64	3·93	-0·67	0·502	-10·34	5·06
Intervention status						
1	2·52	2·74	0·92	0·356	-2·84	7·88

Model Adjusted QPR means

	Margin	Std.Err.	[95%Conf.	Interval]
Sex				
Male	55·00	0·83	53·34	56·61
Female	52·32	1·65	49·08	55·56
Age group				
17-30 years	53·84	1·64	50·62	57·07
30-49 years	54·82	1·18	52·51	57·13
50 years and over	50·83	1·61	47·68	53·99
Timepoint				
T0	54·60	1·39	51·88	57·34
T1	53·72	1·51	50·76	56·68
T2	51·96	2·79	46·49	57·43
Intervention status or sector				
0	52·25	2·18	47·97	56·54
1	54·78	0·96	52·90	56·65

Model 2.3 Stream-two QPR mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=265.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·68	1·94	-1·38	·168	-6·49	1·13

Age group						
30-49 years	·99	1·78	·55	·58	-2·50	4·48
50 years and over	-2·99	2·13	-1·41	·16	-7·16	1·18
Timepoint						
T1	-·88	2·80	-·32	·75	-6·37	4·61
T2	-2·67	4·15	-·64	·52	-10·81	5·47
Intervention status						
1	2·54	2·88	·88	·37	-3·09	8·18
Sector						
2	·23	2·07	·11	·91	-3·83	4·28

Model Adjusted QPR means

	QPR mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	55·00	·81	53·40	56·57
Female	52·31	1·74	49·00	55·71
Age group				
17-30 years	53·83	1·67	50·55	57·11
30-49 years	54·82	1·20	52·47	57·17
50 years and over	50·84	1·59	47·73	53·95
Timepoint				
T0	54·61	1·45	51·77	57·45
T1	53·73	1·49	50·81	56·64
T2	52·00	3·00	46·19	57·60
Intervention status				
0	52·24	2·27	48·00	56·69
1	54·79	·97	53·00	56·70
Sector				
1	53·38	1·82	49·81	57·00
2	53·61	·67	53·00	54·92

Model 2.4 Stream-two QPR mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.
Number of obs=265.

Step variable omitted due to collinearity == model the same as 2.3

Appendix 11

Stream-two models of the secondary outcomes
Model 3.1 Stream-two Warrick mixed model with fixed factors of time and intervention status, and clusters as random.
Number of obs=272.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-·94	2·26	-·41	·68	-5·37	3·50
T2	-1·84	3·85	-·48	·63	-9·39	5·71
Intervention status						
1	1·99	2·42	·82	·41	-2·76	6·74

Model Adjusted means

	Warrick mean	Std.Err.	[95%Conf.	Interval]
Timepoint				
T0	42·84	1·40	40·08	45·59
T1	41·89729	1·22	39·50	44·30
T2	41·00	2·77	35·57	46·43
Intervention status				
0	40·92	1·97	37·07	44·78
1	42·91	·99	40·97	44·85

Model 3.2 Stream-two Warrick mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.
Number of obs=268.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·08	1·271	-1·64	·10	-4·57	·41
Age group						
30-49 years	·80	1·14	·71	·48	-1·43	3·03
50 years and over	-·32	1·13	-·28	·78	-2·53	1·89
Timepoint						

T1	-1·50	2·41	-·62	·53	-6·23	3·23
T2	-2·38	4·04	-·59	·56	-10·31	5·55
Intervention status						
1	2·36	2·48	·95	·34	-2·50	7·22

Model Adjusted means

	Warrick mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	43·05	1·05	41·00	45·10
Female	40·97	1·28	38·47	43·47
Age Group				
17-30 years	41·58	1·43	38·77	44·39
30-49 years	42·38	1·21	40·01	44·76
50 years and over	41·26	1·02	39·26	43·26
Timepoint				
T0	43·23	1·51	40·27	46·18
T1	41·73	1·29	39·21	44·25
T2	40·85	2·84	35·282	46·42
Intervention status				
0	40·73	2·00	36·82	44·64
1	43·092	1·01	41·11	45·07

Model 3.3 Stream-two Warrick mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=268.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·1	1·33	-1·58	·11	-4·71	·51
Age group						
30-49 years	·815	1·13	·72	·47	-1·39	3·02
50 years and over	-·30	1·17	-·25	·80	-2·60	2·00
Timepoint						
T1	-1·51	2·46	-·61	·54	-6·33	3·31
T2	-2·42	4·15	-·58	·56	-10·55	5·72
Intervention status						
1	2·39	2·57	·93	·35	-2·66	7·43
Sector						
2	·33	2·03	·16	·87	-3·65	4·31

Model Adjusted means

	Warrick mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	43·06	1·03	41·04	45·07
Female	40·96	1·32	38·37	43·54
Age group				
17-30 years	41·57	1·46	38·72	44·41
30-49 years	42·38	1·22	39·99	44·78
50 years and over	41·27	1·02	39·28	43·26
Timepoint				
T0	43·24	1·55	40·21	46·28
T1	41·73	1·28	39·23	44·24
T2	40·82	2·90	35·14	46·51
Intervention status				
0	40·72	2·05	36·71	44·73
1	43·10	1·03	41·09	45·12
Sector				
1	41·74	1·61	38·59	44·89
2	42·07	1·12	39·87	44·27

Model 3.4 Stream-two Warrick mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.

Number of obs=268.

Step omitted due to collinearity == model the same as 3.3.

Model 4.1 Stream-two INSPIRE_S_score mixed model with fixed factors of time and intervention status, and clusters as random.

Number of obs=251.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-3·09	4·93	-·63	·53	-12·76	6·57

T2	-2·65	6·22	-·43	·67	-14·84	9·53
Intervention status						
1	1·23	5·45	·23	·82	-9·44	11·91

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
Timepoint				
T0	64·47	3·64	57·33	71·61
T1	61·38	2·74	56·00	66·76
T2	61·82	3·10	55·74	67·90
Intervention status				
0	61·74	3·32	55·24	68·25
1	62·98	2·86	57·36	68·59

Model 4.2 Stream-two INSPIRE_S_score mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs=248.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-2·61	3·19	-·819	·41	-8·86	3·64
Age group						
30-49 years	2·19	5·41	·40	·69	-8·41	12·78
50 years and over	1·30	4·89	·26	·79	-8·29	10·89
Timepoint						
T1	-3·79	4·67	-·81	·42	-12·94	5·36
T2	-3·01	6·06	-·50	·62	-14·89	8·86
Intervention status						
1	1·29	5·36	·24	·81	-9·21	11·80

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	63·88	2·64	58·72	69·05
Female	61·27	1·81	57·72	64·83
Age group				
17-30 years	60·97	4·26	52·62	69·31
30-49 years	63·15	2·30	58·65	67·66
50 years and over	62·26	2·10	58·15	66·38
Timepoint				
T0	64·90	3·53	57·99	71·83
T1	61·12	2·72	55·78	66·45
T2	61·89	3·07	55·88	67·90
Intervention status				
0	61·75	3·28	55·32	68·17
1	63·04	2·87	57·41	68·67

Model 4.3 Stream-two INSPIRE_S_score mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.

Number of obs=248.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-3·49	3·03	-1·15	·25	-9·43	2·46
Age group						
30-49 years	2·43	5·19	·47	·64	-7·74	12·61
50 years and over	1·63	4·82	·34	·73	-7·82	11·09
Timepoint						
T1	-4·34	3·44	-1·26	·207	-11·08	2·40
T2	-4·21	4·50	-·93	·35	-13·04	4·62
Intervention status						
1	2·03	4·46	·45	·65	-6·72	10·78
Sector						
2	7·55	2·50	3·02	·00	2·66	12·45

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
Sex				
Male	64·37	2·19	60·09	68·66
Female	60·89	1·68	57·60	64·17
Age group				

Model Adjusted means

	INSPIRE_S mean	Std.Err.	[95%Conf.	Interval]
17-30 years	60·74	4·02	52·86	68·63
30-49 years	63·18	2·09	59·24	67·11
50 years and over	62·38	2·02	58·43	66·33
Timepoint				
T0	65·46	2·32	60·91	70·01
T1	61·12	2·34	56·53	65·71
T2	61·26	2·68	56·00	66·50
Intervention status				
0	61·37	2·54	56·40	66·35
1	63·40	2·51	58·47	68·33
sector				
1	59·01	1·21	56·65	61·38
2	66·57	2·19	62·27	70·86

Model 4.4 Stream-two INSPIRE_R_score mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.

Number of obs=268.

Step omitted due to collinearity == model the same as 4.3.

Model 5.1 Stream-two INSPIRE_R_score mixed model with fixed factors of time and intervention status, and clusters as random.

Number of obs=263.

	b	se	z	p value	[95%Conf.	Interval]
Timepoint						
T1	-1·72	4·09	-·42	·67	-9·74	6·30
T2	1·41	5·84	·24	·81	-10·04	12·86
Intervention status						
1	2·82	4·87	·58	·56	-6·72	12·36

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
Timepoint				
T0	74·14	3·43	21·62	80·87
T1	72·43	2·44	29·70	77·21
T2	75·55	3·06	24·73	81·54
Intervention status				
0	72·31	2·75	26·27	77·71
1	75·14	2·89	25·99	80·80

Model 5.2 Stream-two INSPIRE_R_score mixed model with fixed factors of sex, age-group, time and intervention status, and clusters as random.

Number of obs=259.

	b	se	z	pvalue	[95%Conf.	Interval]
Sex						
Female	·78	3·10	·25	·80	-5·30	6·86
Age group						
30-49 years	2·36	5·32	·44	·66	-8·06	12·78
50 years and over	4·26	5·45	·78	·43	-6·42	14·95
Timepoint						
T1	-2·02	4·19	-·48	·63	-10·22	6·19
T2	1·10	5·74	·19	·85	-10·16	12·36
Intervention status						
1	2·45	4·86	·50	·61	-7·07	11·97

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
Sex				
Male	73·27	2·30	68·77	77·76
Female	74·05	1·93	70·27	77·83
Age Group				
17-30 years	71·35	4·41	62·70	80·00
30-49 years	73·71	2·30	69·20	78·21
50 years and over	75·61	2·03	71·63	79·59
Timepoint				

T0	74·35	3·33	67·82	80·88	
T1	72·33	2·44	67·54	77·12	
T2	75·45	3·12	69·33	81·57	
Intervention status					
0	72·49	2·84	66·91	78·06	
1	74·94	2·79	69·47	80·40	

Model 5.3 Stream-two INSPIRE_R_score mixed model with fixed factors of sector, sex, age-group, time and intervention status, and clusters as random.
Number of obs=259.

	b	se	z	p value	[95%Conf.	Interval]
Sex						
Female	-·20	2·86	-·072	·94	-5·81	5·40
Age group						
30-49 years	2·47	4·89	·50	·61	-7·11	12·06
50 years and over	4·74	5·29	·90	·37	-5·63	15·1
Timepoint						
T1	-2·51	2·99	-·84	·40	-8·38	3·35
T2	-·44	3·43	-·13	·90	-7·16	6·27
Intervention status						
1	3·28	3·41	·96	·33	-3·39	9·97
Sector						
2	8·22	1·71	4·8	0	4·87	11·58

Model Adjusted means

	INSPIRE_R mean	Std. Err.	[95%Conf.	Interval]
Sex				
Male	73·76	1·73	42·59	77·16
Female	73·56	1·62	45·40	76·73
Age group				
17-30 years	71·12	4·03	17·66	79·02
30-49 years	73·60	1·76	41·78	77·05
50 years and over	75·86	2·06	36·87	79·90
Timepoint				
T0	74·95	1·87	40·04	78·62
T1	72·44	1·75	41·42	75·87
T2	74·51	2·46	30·25	79·34
Intervention status				
0	72·04	1·97	36·65	75·89
1	75·33	1·86	40·45	78·98
sector				
1	70·06	·98	71·53	71·99
2	78·29	1·47	53·12	81·18

Model 5.4 Stream-two INSPIRE_R_score mixed model with fixed factors of sector, step-group, sex, age-group, time and intervention status, and clusters as random.
Number of obs=259

Step omitted due to collinearity == model the same as 5.3

Appendix 12

Interaction term model for stream-one QPR outcome but with reference category as the first group
Below are identical models as shown in Appendix 8. However, the below models have the first group displaying as the reference group.
The context was one of typically declining QPR scores in the groups not receiving the intervention, possibly for reasons of organisational context discussed in the manuscript. While formally speaking, chose of reference categories does not affect the outcomes when presented as model adjusted means, varying the reference categories is a way to enable inspection of differences between categories. In the context of the step-wedge design and with declining QPR scores in non-intervention groups, that T1 non-intervention scores can be seen as most indicative of a 'baseline' and this guided chose of reference category for inclusion in the manuscript. For completeness we include another set of model presentations base on T0 reference categories here.

Model 1.5b. Stream-one QPR has interaction item of time and intervention status, and cluster as random.
Number of obs=942

Timepoint	Intervention	Coefficient	Robust Std Err.	P>|z|	95% CI	
T0	No	Reference				
T1	No	-3·59	1·13	0·002	-5·81	-1·37
T1	Yes	0·56	1·24	0·651	-1·88	3·00
T2	Yes	-0·63	1·24	0·618	-3·10	1·84

Model 1.6b. Stream-one QPR has interaction item of time and intervention status, and fixed variables of age-group (<30; 30-49; 50 years and over) and sex (Male/Female) and cluster as random.

Number of obs=942.

		Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex						
	Male			Reference		
	Female	-0·86	1·05	0·41	-2·92	1·20
Age Category						
	17-29			Reference		
	30-49	-0·88	0·86	0·31	-2·56	0·81
	50-75	-3·39	0·92	<0·001	-5·21	-1·58
Timepoint	**Intervention**					
T0	No			Reference		
T1	No	-3·20	1·07	0·003	-5·29	-1·10
T1	Yes	0·54	1·26	0·669	-1·93	3·02
T2	Yes	-0·46	1·33	0·688	-2·68	1·77

Model 1.7b. Stream-one QPR. Model has same variables as Model 1.6b plus sector (PMHS; MHCSS) as fixed.

Number of obs=942.

			Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex							
		Male (reference)					
		Female	-0·76	1·05	0·47	-2·81	1·29
Age Category							
		17-29 (reference)					
		30-49	-0·90	0·84	0·28	-2·55	0·75
		50-75	-3·37	0·91	0·00	-5·15	-1·59
Timepoint	**Intervention**	**Sector**					
T0	No	**PMHS**	reference				
T0	No	**MHCSS**	-0·97	2·15	0·65	-5·19	3·24
T1	No	**PMHS**	-3·56	1·47	0·02	6·43	-0·68
T1	No	**MHCSS**	-3·53	2·83	0·21	-9·07	2·01
T1	Yes	**PMHS**	0·44	1·89	0·87	-3·26	4·14
T1	Yes	**MHCSS**	-0·23	2·14	0·92	-4·01	3·95
T2	Yes	**PMHS**	0·74	1·50	0·62	-2·18	3·67
T2	Yes	**MHCSS**	-3·24	2·43	0·18	-8·01	1·53

Model 1.8b. Stream-one QPR. Model has same variables as model 1.7b plus 'step' group as fixed variable. Model also examined interactions between four variables (sector, step, time and intervention).

Number of obs=942.

				Coefficient	Robust Std Err.	P>|z|	95% CI	
Sex								
			Male (reference)					
			Female	-0·76	1·06	0·47	-2·85	1·32
Age Category								
			17-29 (reference)					
			30-49	-0·87	0·87	0·32	-2·56	0·83
			50-75	-3·37	0·94	0·00	-5·22	-1·52
Timepoint	**Intervention**	**Step Group**	**Sector**					
T0	No	1	PMHS	reference				
T0	No	2	PMHS	-3·47	3·50	0·33	-10·38	3·46
T1	No	2	PMHS	-5·83	3·40	0·09	-12·49	0·84
T1	Yes	1	PMHS	-0·81	1·70	0·63	-4·14	2·53
T2	Yes	1	PMHS	-1·35	1·76	0·44	-4·81	2·11
T2	Yes	2	PMHS	-0·91	2·11	0·67	-5·05	3·24
T0	No	1	MHCSS	-3·84	1·75	0·03	-7·27	-0·41
T0	No	2	MHCSS	-1·68	1·82	0·36	-5·27	1·90

T1	No	2	MHCSS	-4·54	2·94	0·12	-10·31	1·23
T1	Yes	1	MHCSS	-2·72	1·59	0·09	-5·84	0·40
T2	Yes	1	MHCSS	-6·12	1·42	<0·001	-8·89	-3·34
T2	Yes	2	MHCSS	-3·89	3·09	0·21	-9·95	2·16

References

1. Victorian Healthcare Association. State budget submission 2015-2016. Melbourne: Victorian Healthcare Association 2015.

2. Victorian Healthcare Association. State budget submission 2017-2018. Melbourne: Victorian Healthcare Association 2017.

3. Silburn K. Recommissioning community mental health support services and alcohol and other drugs treatment services in Victoria: Report on findings from interviews with senior personnel from both sectors: La Trobe University, Australian Institute for Primary Care and Ageing; 2015.

4. Koop K. NDIS and state mental health reform: opportunities and risks for Victoria. *New Paradigm* 2014 **Summer**.

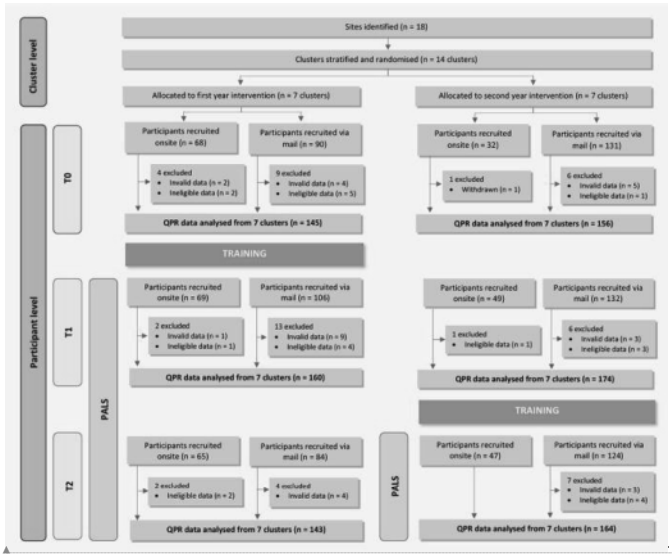
5. Mental Health Act 2014.

6. Klein JH, Young T. Health care: a case of hypercomplexity? *Health Systems* 2015; **4**(2): 104-10.

7. Long KM, McDermott F, Meadows GN. Being pragmatic about healthcare complexity: our experiences applying complexity theory and pragmatism to health services research. *BMC Med* 2018; **16**(1): 94.

8. Long KM, Mc Dermott F, Meadows GN. Mental health services management and simulation modelling: A project evaluation drawing on pragmatism, complexity theory and the critical incident approach. (unpublished).

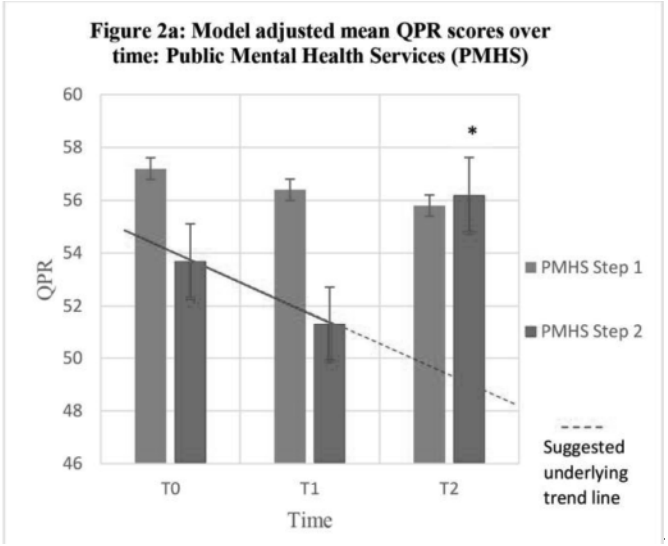
~~Figures~~

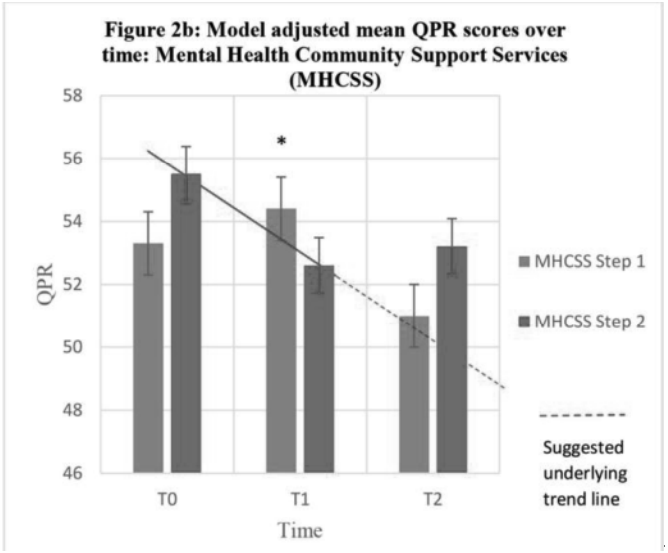


~~**Figure 1 Consort Chart**~~

~~Note. PALS = PULSAR Active Learning Sessions~~

Formatted: Font: 12 pt

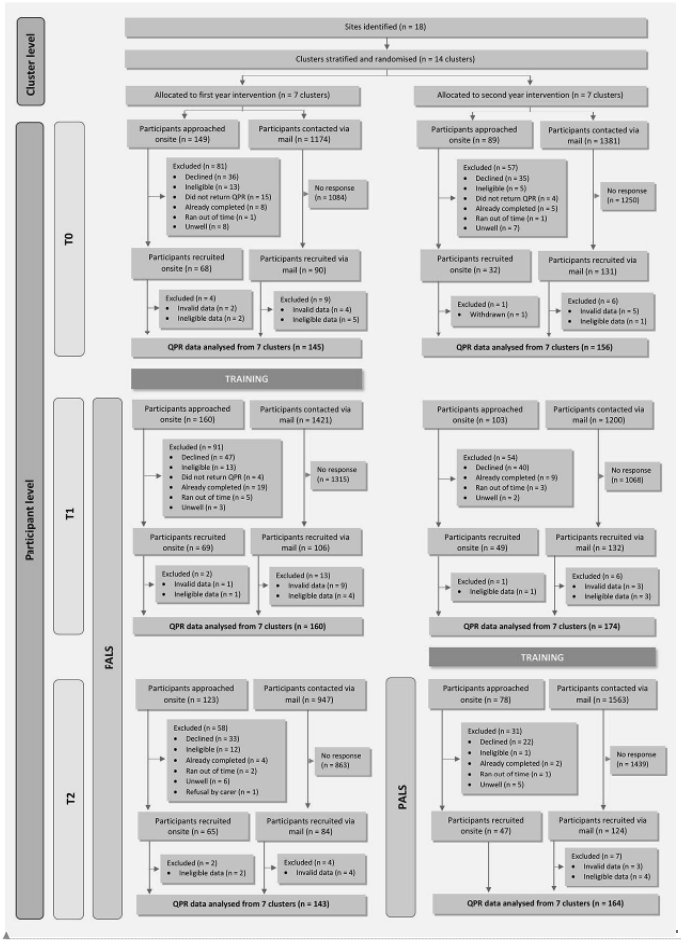




~~**Figure 2 QPR scores by sector over time.**~~

~~*Change p < .01 by pairwise comparison with previous time-point.~~
~~Note. Step-one group (blue) received intervention in year 1. Step-two group (red) received intervention in year 2. A dashed line is plotted by connecting mean scores for Step Group 2 at baseline (T0) and year 1 (T1). This dashed line between two control periods suggests a decreasing underlying trend in both sectors.~~

~~Supplementary figure~~



~~Figure S1: Detailed CONSORT chart~~

~~Note. PALS = PULSAR Active Learning Sessions~~

Formatted: Font: (Default) +Headings, 16 pt, Bold

References

1. Deegan P. Recovery as a journey of the heart. *Psychiatric Rehabilitation Journal* 1996; **19**(3): 91-7.

2. Slade M, Bird V, Le Boutillier C, et al. Development of the REFOCUS intervention to increase mental health team support for personal recovery. *Br J Psychiatry* 2015; **207**(6): 544-50.

3. Shepherd G, Boardman J, Slade M. Making recovery a reality. London: Sainsbury Centre for Mental Health 2008.

4. World Health Organization. Mental health action plan 2013 - 2020. Geneva: WHO Document Production Services; 2013.

5. Australian Government Department of Health. National mental health policy 2008. Canberra: Commonwealth of Australia 2009.

6. Slade M, Bird V, Le Boutillier C, Williams J, McCrone P, Leamy M. REFOCUS Trial: protocol for a cluster randomised controlled trial of a pro-recovery intervention within community based mental health teams. *BMC Psychiatry* 2011; **11**(1): 185.

7. Slade M, Bird V, Clarke E, et al. Supporting recovery in patients with psychosis through care by community-based adult mental health teams (REFOCUS): a multisite, cluster, randomised, controlled trial. *Lancet Psychiatry* 2015; **2**(6): 503-14.

8. Ajzen I. The theory of planned behaviour: reactions and reflections. Psychol Health 2011; **26**: 1113-27.

9. Le Boutillier C, Chevalier A, Lawrence V, et al. Staff understanding of recovery-orientated mental health practice: a systematic review and narrative synthesis. *Implement Sci* 2015; **10**(1): 1-14.

10. Neil ST, Kilbride M, Pitt L, et al. The questionnaire about the process of recovery (QPR): A measurement tool developed in collaboration with service users. *Psychosis* 2009; **1**(2): 145-55.

11. Bird V, Leamy M, Le Boutillier C, Williams J, Slade M. REFOCUS: Promoting Recovery in Mental Health Services. 2nd ed. London: Institute of Psychiatry, King's College London; 2014.

12. Shawyer F, Enticott JC, Brophy L, et al. The PULSAR Specialist Care protocol: a stepped-wedge cluster randomized control trial a training intervention for community mental health teams in recovery-oriented practice. *BMC Psychiatry* 2017; **17**: 172.

13. Victorian Healthcare Association. State budget submission 2015-2016. Melbourne: Victorian Healthcare Association 2015.

14. Victorian Healthcare Association. State budget submission 2017-2018. Melbourne: Victorian Healthcare Association 2017.

15. Chief Parliamentary Counsel. Mental Health Act 2014 In: Government of Victoria, editor. No 26 of 2014 2014.

16. Silburn K. Recommissioning community mental health support services and alcohol and other drugs treatment services in Victoria: Report on findings from interviews with senior personnel from both sectors: La Trobe University, Australian Institute for Primary Care and Ageing; 2015.

17. Bonyhady B. Bruce Bonyhady Speech: 'Reducing the Inequality of Luck'. 2015. *https://www.ndis.gov.au/chair-speech-reducing-inequality* (accessed 01.05.2018.

18. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions an introduction to the new Medical Research Council guidance. *Evidence-Based Public Health: Effectiveness and Efficiency* 2010: 185-203.

19. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ quality & safety* 2014; **23**: 290-8.

20. PULSAR-REFOCUS team. PULSAR Manual: Recovery-promoting Relationships and Working Practices for Specialist and Community Mental Health Services (or Secondary Care Staff). 2nd ed. Melbourne: Monash University; 2016.

21. Australian Government. National Standards for Mental Health Services 2010. Australia: Commonwealth of Australia; 2010.

22. Williams J, Leamy M, Pesola F, Bird V, Le Boutillier C, Slade M. Psychometric evaluation of the Questionnaire about the Process of Recovery (QPR). *Br J Psychiatry* 2015; **207**(6): 551-5.

23. Law H, Neil ST, Dunn G, Morrison AP. Psychometric properties of the questionnaire about the process of recovery (QPR). *Schizophrenia research* 2014; **156**(2-3): 184-9.

24. Statacorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.

25. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012; **345**.

26. Torres-Reyna O. Multilevel Analysis (ver. 1.10). Princeton University Library, Data and Statistical Services; 2015.

27. Selya AS, Rose JS, Dierker LC, Hedeker D, Mermelstein RJ. A practical guide to calculating Cohen's f(2), a measure of local effect size, from PROC MIXED. *Froniers in Psychology* 2012.

28. Brysbaert M, Stevens M. Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition* 2018; **1**(1).

29. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008; **337**.

30. Sharkey K, Savulescu J, Aranda S, Schofield P. Clinician gate-keeping in clinical research is not ethically defensible: an analysis. *Journal of medical ethics* 2010; **36**(6): 363-6.

*Protocol

Shawyer et al. BMC Psychiatry (2017) 17:172
DOI 10.1186/s12888-017-1321-3

BMC Psychiatry

STUDY PROTOCOL Open Access

The PULSAR Specialist Care protocol: a stepped-wedge cluster randomized control trial of a training intervention for community mental health teams in recovery-oriented practice

Frances Shawyer[1*], Joanne C. Enticott[1,2], Lisa Brophy[3,4], Annie Bruxner[1], Ellie Fossey[5], Brett Inder[6], John Julian[1], Ritsuko Kakuma[4], Penelope Weller[7], Elisabeth Wilson-Evered[8], Vrinda Edan[1], Mike Slade[9] and Graham N. Meadows[1,2,4]

Abstract

Background: Recovery features strongly in Australian mental health policy; however, evidence is limited for the efficacy of recovery-oriented practice at the service level. This paper describes the Principles Unite Local Services Assisting Recovery (PULSAR) Specialist Care trial protocol for a recovery-oriented practice training intervention delivered to specialist mental health services staff. The primary aim is to evaluate whether adult consumers accessing services where staff have received the intervention report superior recovery outcomes compared to adult consumers accessing services where staff have not yet received the intervention. A qualitative sub-study aims to examine staff and consumer views on implementing recovery-oriented practice. A process evaluation sub-study aims to articulate important explanatory variables affecting the interventions rollout and outcomes.

Methods: The mixed methods design incorporates a two-step stepped-wedge cluster randomized controlled trial (cRCT) examining cross-sectional data from three phases, and nested qualitative and process evaluation sub-studies. Participating specialist mental health care services in Melbourne, Victoria are divided into 14 clusters with half randomly allocated to receive the staff training in year one and half in year two. Research participants are consumers aged 18–75 years who attended the cluster within a previous three-month period either at baseline, 12 (step 1) or 24 months (step 2). In the two nested sub-studies, participation extends to cluster staff. The primary outcome is the Questionnaire about the Process of Recovery collected from 756 consumers (252 each at baseline, step 1, step 2). Secondary and other outcomes measuring well-being, service satisfaction and health economic impact are collected from a subset of 252 consumers (63 at baseline; 126 at step 1; 63 at step 2) via interviews. Interview-based longitudinal data are also collected 12 months apart from 88 consumers with a psychotic disorder diagnosis (44 at baseline, step 1; 44 at step 1, step 2). cRCT data will be analyzed using multilevel mixed-effects modelling to account for clustering and some repeated measures, supplemented by thematic analysis of qualitative interview data. The process evaluation will draw on qualitative, quantitative and documentary data.

(Continued on next page)

* Correspondence: frances.shawyer@monash.edu
[1]Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Dandenong Hospital, 126 - 128 Cleeland St, Dandenong, Victoria 3175, Australia
Full list of author information is available at the end of the article

(Continued from previous page)

Discussion: Findings will provide an evidence-base for the continued transformation of Australian mental health service frameworks toward recovery.

Trial Registration: Australian and New Zealand Clinical Trial Registry: ACTRN12614000957695. Date registered: 8 September 2014.

Keywords: Recovery, Recovery-oriented Practice, Specialist Mental Health Services, Mental Health, Co-production, Co-design, Training, Psychiatry, Randomized Controlled Trial (RCT), Complex Intervention

Background

Recovery-oriented practice involves facilitating a process of change through which individuals who have been diagnosed with mental illness are supported to reclaim their personal identity, live a self-directed life, and strive to reach their full potential [1, 2]. This can be seen as a paradigm shift in specialist mental health service delivery, from a focus on ameliorating symptoms to an approach that recognises people's strengths, self-capacity and potential for personal recovery, even in the context of ongoing symptoms or disability [3, 4]. The history of the international recovery movement is longstanding and influenced by the consumer movement as well as emerging evidence that challenges more pessimistic assumptions about recovery from severe and persistent mental illness [5, 6]. Note that in this paper we will use the term consumer to refer to a person with a diagnosis of mental illness or who uses mental health services. The impact of this paradigm shift towards recovery can be identified in mental health policy, practice and law in all Australian states and territories, especially in the last 10 years [7], and is gradually transforming services.

The development of a recovery orientation in mental health in Victoria

Since the concept of recovery first emerged from the consumer movement in the 1970s and 1980s, the reorientation of mental health policy and services toward recovery has gained increasing momentum in the Victorian mental health sector [8]. At the national level, recovery was first formally endorsed in the 2003–2008 Australian National Mental Health Plan [9]. Subsequent developments in the community-managed health sector accelerated the Australian recovery movement over the following decade, including the establishment of early intervention alternatives to inpatient treatment such as the sub-acute Prevention and Recovery Care (PARC) programs. In 2011, the Victorian Government commissioned a framework document supporting the development of evidence-based recovery-oriented mental health services with an emphasis on facilitating personal recovery and dismantling barriers to full participation in community life for people with experiences of mental illness [8]. This was followed in 2014 by the implementation of a new Mental Health Act in Victoria that established recovery as a fundamental guiding principle in the provision of mental health care. Recovery has thus emerged as a core feature of contemporary reform to mental health service planning and delivery at both the state and national level. Complementing these developments has been an increasing emphasis on the importance of "co-design" or "co-production" to ensure that consumers, families and carers are centrally involved in the design, development and delivery of mental health services [10, 11]. Despite these reforms, the mental health care system still has a long way to go in being responsive to the cultural and linguistic diversity of the Australian population [12]. Little is known about the effectiveness of mental health interventions across people of different cultural and linguistic groups and whether the contemporary emphasis on recovery orientated practice is having the presumed positive impacts on consumers.

REFOCUS

The value and efficacy of system-wide transformation to focus on recovery is yet to be empirically established in Australia. In the UK, a staff training intervention (called REFOCUS) promoting personal recovery and enabling organizational change in specialist mental health services has been developed and trialled [13]. Based on a systematic review and narrative synthesis of existing literature on recovery, the REFOCUS team developed a conceptual framework of personal recovery, which identified five key recovery processes denoted by the acronym "CHIME": Connectedness, Hope, Identity, Meaning, and Empowerment [14]. This conceptual framework informed the development of a team-based training intervention for community mental health teams in England that was designed to promote recovery through changes in staff and team skills, knowledge, behaviour, values and relationships with consumers [13]. In a large scale cluster randomized controlled trial (cRCT), the outcomes of usual care plus the REFOCUS intervention were compared with usual care only (control) in 27 community mental health teams delivering services to adult patients with psychotic disorders. The primary outcome,

personal recovery as assessed using the Questionnaire about the Process of Recovery (QPR), did not differ between the REFOCUS intervention group and controls, although staff-rated functioning and unmet needs did improve in the intervention group [15]. The authors suggest that implementation was the central challenge, and when high-participating teams were compared with low-participating teams, higher participation was associated with higher staff-reported recovery-promotion behaviour and improved consumer-rated QPR. Challenges in implementing the intervention at the team level included variability in staff participation and adherence to recovery-oriented training procedures [16] along with the diluting effects of staff turnover. Participant attrition was higher than anticipated (26% vs 7%) resulting in a reduction in planned statistical power. A further proposed possible reason for the overall lack of difference between the intervention and control group on recovery outcomes is that the 12-month timeframe may have been of insufficient length for the intervention to take effect. On average, patient participants had been using mental health services for more than 15 years, suggesting the possibility of established staff-consumer relationships and entrenched ways of relating to services and problems that may take longer than 1 year to change [17].

PULSAR

In the Principles Unite Local Services Assisting Recovery (PULSAR) Specialist Care trial, REFOCUS training materials and the research design have been adapted to enable the testing of the intervention in specialist mental health care services in Australia. The PULSAR staff training intervention aims to train community mental health staff in recovery-oriented practice, so as to embed recovery principles in mental health service delivery in the southern metropolitan region of Victoria, Australia. The PULSAR Specialist Care trial is part of the broader PULSAR research program focused on promoting recovery-oriented practices which also includes the Australian primary care sector [18]. The study components were co-designed with a consumer academic (VE), and the involvement of PULSAR Lived Experience Advisory Panel (LEAP) created for the project (see *Leadership structure* below) and facilitated by VE. This paper outlines the PULSAR Specialist Care study protocol. The protocol follows the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines [19].

Objectives

Using a mixed methods design, the primary objective of the PULSAR Specialist Care study is to evaluate whether adults accessing study cluster specialist mental health services where staff receive the recovery-oriented practice training intervention report superior recovery outcomes compared to adults accessing services where staff have not received the intervention. The following research questions will be addressed:

1. From pre- to post-intervention, do consumers in intervention clusters report greater improvements in a) personal recovery b) health and well-being, and c) perceived need and satisfaction with services compared with consumers receiving care in control groups?
2. From pre- to post-intervention, do ethnic minority consumers in intervention clusters report greater improvements on measures of personal recovery compared to ethnic minority consumers receiving care during control phases?

A nested qualitative sub-study involving consumers and staff will be conducted. For consumers, the research question that will be addressed is:

1. How do consumers experience and view the support for their recovery in services where the PULSAR training has taken place?

For staff, the research questions to be considered are:

1. What factors help and hinder working in a recovery-oriented manner, from the perspective of staff who have received the PULSAR training intervention?
2. What experiences and dilemmas are encountered when implementing recovery-oriented practices within different parts of the Australia mental health service system, and what strategies are used to address the issues identified?

A nested process evaluation sub-study aims to examine quantitative and qualitative data including documents and processes related to training implementation and the uptake of new ways of working in order to articulate important explanatory variables relating to clusters that affected the rollout of the intervention and potentially influenced the study outcomes.

Methods

Overall design

The PULSAR Specialist Care project is one of two multisite two-step stepped-wedge cRCTs within the broader PULSAR research program [18]. The study design of the PULSAR Specialist Care project is a mixed methods design incorporating a two-step stepped-wedge cluster randomized controlled trial (cRCT) examining cross-sectional data from three phases, and nested qualitative and process evaluation sub-studies (see Fig. 1 and

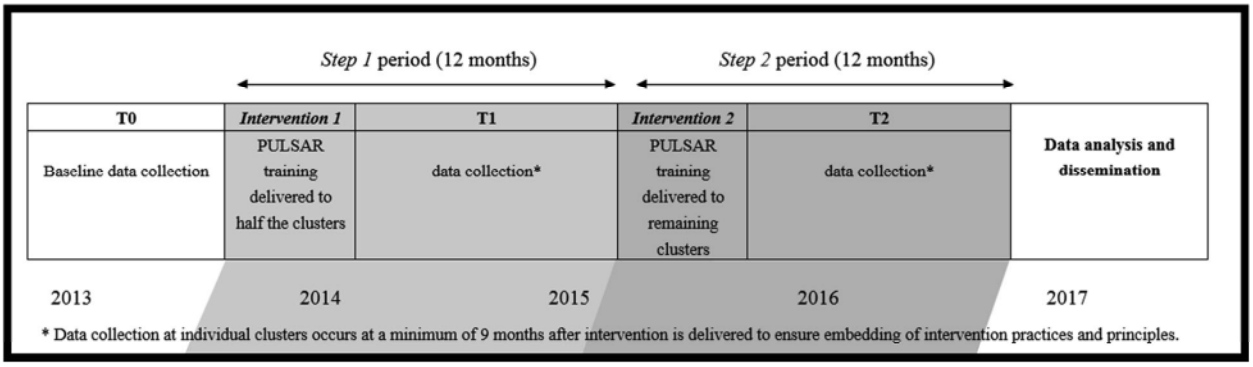


Fig. 1 PULSAR Specialist Care study design and planned timelineNotes. Indicated at the *bottom* is the study year. The two-step cluster randomise control trial is shown: half the clusters receive the intervention in step 1 and the remaining clusters receive the intervention in step 2. For a summary of the study evaluations see Table 1

Table 1). Co-design with a consumer academic who is an author (VE) was a key design driver, and this co-design began early when the initial protocols were drafted for the application for funding.

A cluster randomized design was selected to minimize the threat of contamination between treatment and control groups, as the intervention is administered at the service (cluster) level [15, 20], these clusters being specialist mental health care services in Melbourne, Victoria (see Table 2). The stepped-wedge design was implemented for pragmatic and ethical reasons; in addition to allowing the intervention to be staggered across study clusters, the stepped design offers the ethical advantage of enabling all participating clusters to receive an intervention that is predicted to be beneficial [21]. A mixed method design was utilized for reasons including that the integration of quantitative and qualitative methods provides a richer dataset than can be gained with either approach alone [22, 23]. This in turn allows for a comprehensive and multifaceted evaluation of the project outcomes, as well as further explaining those outcomes and enhancing the project's credibility. The latter is an important consideration in large-scale projects involving multiple stakeholders, as is the case with the PULSAR project [23]. There are additional benefits to using qualitative and quantitative methods at different stages in the intervention trial; for example, exploratory qualitative techniques were initially used to identify potential obstacles to implementing recovery-oriented practice in service settings and to inform the intervention design, while quantitative methods are used to evaluate the effectiveness of the training intervention. Furthermore, following the intervention trial, qualitative methods are being used to gain a more nuanced understanding of staff experiences of implementing recovery-oriented practices and consumer views of these practices across different settings within Victoria's mental health service system.

The PULSAR Specialist Care trial will take 4 years to complete, beginning in 2013 and concluding in 2017, see Fig. 1. The original study protocol (documented in the Australian New Zealand Clinical Trials Registry, or ANZCTR) was developed over a period of 18 months in consultation with Chief Investigators (CIs) and an advisory committee comprised of representatives from local specialist and community care organizations, consumers and family/carers, and experts in legal, cultural and educational aspects of mental health service delivery in Victoria. Although minor adaptations have been made to the original protocol, as outlined below, the over-arching two-step stepped-wedge cRCT design remains unchanged and the trial is on target to reach completion within the anticipated 4 year timeframe. All adaptations to the study protocol were considered by, and required the approval of, the appropriate Module Committee governing the relevant aspect of the project (see *Study leadership* section) along with the governing Human Research Ethics Committees.

The PULSAR training intervention is being delivered to 14 specialist mental health care clusters (see Table 2), with clusters randomized to receive the intervention 12 months apart, as shown in Fig. 1. To ensure that cluster types are balanced, stratified randomization was applied, for the strata see Table 2.

Explanation for choice of comparators

The design was developed to combine the rigor of a cluster randomized trial with the pragmatic approach of the stepped wedge design to implement and evaluate the intervention at all sites [20, 21, 24]. Control sites are those that are yet to receive the intervention. Since all sites eventually receive the intervention, data from sites in control phases will be compared with data from sites that have received the intervention.

Table 1 Evaluation plan for the PULSAR Specialist Care study

	Sub-study name	Evaluation design	Unit of analysis	Number	Number at each time point	Number in each cluster at each time point	Detectable differences		
							Primary outcome, QPR	Secondary outcome, WEMWBS	Secondary outcome, INSPIRE
cRCT (quantitative data)	Stream 1 (primary analysis)	cross-sectional cRCT (complete step-wedge)	Consumers (mail-out)	756	252 at baseline 252 at step 1 252 at step 2	18	6.34 (medium effect)	NA	NA
	Stream 2	pre- and post-intervention (incomplete step-wedge)	Consumers (interviews)	252 (stream 1 subset)	63 at baseline 126 at step 1 63 at step 2	9	7.68 (medium effect)	4.80 (medium effect)	7.72 (medium effect)
	Stream 3	longitudinal, (same participant, 12-mths apart pre- and post-intervention)	Consumers with diagnosis of psychosis (interviews)	88 (stream 2 subset)	44 at baseline & step 1 44 at step 1 & step 2	6–7	10.94 (medium-large effect)	6.84 (medium-large effect)	11.28 (medium-large effect)
Nested Qualitative study (qualitative data)			Consumers Staff	20–24 20–24	10–12 at step 1 10–12 at step 2	Nested sub-study examining qualitative data collected in study interviews and focus groups			
Nested Process evaluation (both quantitative & qualitative data)			Consumers & staff	The process evaluation assesses a specific set of qualitative, quantitative and documentary data relating to each cluster.					

Notes. The primary analysis examines the primary outcome– the Questionnaire about the Process of Recovery (QPR) - collected in the two-step stepped-wedge cluster randomized controlled trial (cRCT). A subset of consumers in the cRCT participate in study interviews where secondary outcomes measuring well-being, service satisfaction and health economic impact are collected In a yet another subset in the cRCT, longitudinal data are collected via interviews that are 12 months apart from consumers with a diagnosis of a psychotic disorder. Qualitative and process evaluation sub-studies are nested within the overarching cRCT and include information from consumers and staff.

There are no study restrictions on the care provided in control phases. Treatment as usual is described later under the heading "Control".

Study setting and clusters

In Victoria, specialist mental health services include area-based clinical services comprising a range of teams and service types, in particular, inpatient units and community-based continuing care and treatment teams, as well as Mental Health Community Support Services (MHCSS) [25] that provide residential and outreach support. The study setting is the catchment of Monash Health, the largest public health care provider in Victoria, which provides services to a population of over 950,000 in the South-Eastern suburbs of Melbourne, and encompasses a greater population of 1.34 million people [26]. The Monash Health catchment area includes the City of Greater Dandenong, the most culturally diverse municipality in Victoria [27]. Three organisations that operate within the Monash Health catchment are involved in the study including: Monash Health Specialist Clinical Mental Health Services and, from the MHCSS sector, Mind Australia and Ermha.

Fourteen participating specialist mental health care clusters are spread over 18 adult community-based mental health service sites. See Table 2 for a description of study clusters.

Participants

Levels of staff participation

Staff participate in the PULSAR training intervention but no data from individual staff members are collected as part of the cRCT. However, staff are requested to complete a training evaluation at the conclusion of the training and this data will be examined in the process evaluation sub-study. Staff may volunteer for the nested qualitative sub-study.

Staff selection criteria

Staff at the study clusters who receive the PULSAR training intervention must fulfil the following inclusion criteria: (a) working on a part-time or full-time basis within the team in a direct service capacity (and not employed on a casual basis); (b) have an active case load with consumers who are recruited into the evaluation. Staff are ineligible if they are also working in a non-intervention cluster at the time of training (to reduce risk of contamination). Similarly, staff are eligible to participate in the nested qualitative sub-study if currently working at services where the training intervention has been provided; and ineligible if they are either not working at a participating cluster site (even if they have completed the training) or no longer work in a direct service role.

Table 2 Cluster sites and stratification factors

Cluster	Team/service	Organisation	Site	Strata	Team/service description
Cluster 1	Crisis assessment & treatment team	Monash Health	1	A	Crisis Assessment and Treatment Teams (CATTs) provide urgent assessment and short-term treatment to people in psychiatric crisis and play a key role in triaging admissions to hospital.
Cluster 2	Crisis assessment & treatment team	Monash Health	2	A	
Cluster 3	Mobile support and treatment service + Community Care Unit	Monash Health	3 4	B	Mobile Support and Treatment Teams (MSTs) provide intensive long-term support through assertive outreach to people with prolonged and severe mental illness with associated high levels of disability. Community Care Units (CCUs) provide medium to long-term residential rehabilitation in a home-like environment.
Cluster 4	Mobile support and treatment service + Community Care Unit	Monash Health	5 6	B	
Cluster 5	Community Mental Health Service	Monash Health	7	C	Community Mental Health Service (CMHS) provide non-urgent assessment, treatment, case management, and continuing care support to people living in the community over varying periods of time.
Cluster 6	Community Mental Health Service	Monash Health	8	C	
Cluster 7	Continuing Care	Monash Health	9	D	Continuing Care Teams (CCTs) provide non-urgent assessment, treatment, case management, and continuing care support to people living in the community over varying periods of time.
Cluster 8	Continuing Care	Monash Health	10	D	
Cluster 9	PARC – Adult + PARC – Extended	Mind Australia	11 12	E	PARC services (including youth, adult and extended PARCs) provide short-term residential support and treatment to assist in averting acute inpatient admission or facilitate earlier discharge.
Cluster 10	PARC – Youth	Mind Australia	13	E	
Cluster 11	PARC Ermha	Ermha	14	F	
Cluster 12	PARC Ermha	Ermha	15	F	
Cluster 13	Community outreach services	Mind Australia	16	G	Community Outreach Services provide a range of individualized psychosocial support and recovery services.
Cluster 14	Community outreach services site 1 Community outreach services site 2	Ermha	17 18	G	

Note. Clusters are stratified by the team/service type and composition: i.e. there are seven different strata

Levels of consumer participation

In line with the presumption of capacity endorsed by the Victorian *Mental Health Act 2014* and research indicating that participating in research can lead to positive outcomes for people who experience mental health issues [28, 29], the project was designed to provide consumer participants with the opportunity to self-select into multiple levels of involvement. Consumers consenting to participation in the cRCT are offered three levels of involvement, see the three streams outlined in Table 1. In the cRCT stream 1, we collect primary outcome recovery-focused data from consumers recruited cross-sectionally via the mail at baseline (T0), end of year 1 (T1) and end of year 2 (T2), see Fig. 1. In the cRCT stream 2, we additionally collect secondary and other outcome data assessing mental health and wellbeing, service satisfaction, perceived coercion when accessing services, and health economic impact. This information is collected via face-to-face interviews from participants recruited cross-sectionally at pre or post-intervention. Stream 3 consists of a smaller subset of Stream 2 participants who have a clinical diagnosis of psychosis, e.g., schizophrenia, schizoaffective disorder, or bipolar disorder. Participants in this Stream provide longitudinal data via pre- and post-intervention face-to-face interviews. A pool of research assistants trained in study-specific interview procedures and blinded to intervention status, conduct all Stream 2 and 3 interviews. Consumers who volunteer for the nested qualitative substudy do not need to participate in the cRCT. Consumer data from the cRCT and qualitative study may be used in the nested process evaluation.

Consumer selection criteria

Consumers are eligible for participation in any study component if they are receiving care from teams in the participating cluster services. The inclusion criteria for any study involvement are: (a) aged between 18 and 75 years inclusive at time of recruitment; (b) able to provide informed consent; (c) proficient in English; and (d) have accessed a study cluster in the 3 months prior to data collection. An additional inclusion criterion for participation in Stream 3 of the cRCT (see Table 1) is a primary clinical diagnosis of psychosis, e.g., schizophrenia,

schizoaffective disorder, or bipolar disorder, recorded in their medical records. The exclusion criteria for all participants are people who are in prison, people unable to give informed consent, and those unable to speak or read English.

Participant timelines

An overview of the schedule of enrolment, intervention and assessments is shown in Table 3. For staff who participate in the PULSAR training intervention, the overview for the delivery of this training is shown in Fig. 1. In brief, half the study clusters will organise the training to occur with their staff in step 1 and the remaining clusters in step 2. Participating consumers are offered multiple levels of involvement, see above for a summary of the timeline commitments required for the various study involvements (also see Fig. 1 and Table 1). Data collection at individual clusters occurs at a minimum of 9 months after intervention is delivered to ensure embedding of intervention practices and principles.

Intervention

The intervention is a training program delivered to staff teams in participating clusters over two-day workshops (either team or organisational groups) or equivalent hours. The intervention is based on a package of tools promoting recovery-oriented practice in mental health care that were developed by the REFOCUS team [30, 31] and adapted for the Australian public clinical mental health care setting and the MHCSS Sector [25]. The adaptation of the REFOCUS materials was guided by consultations with: the REFOCUS research team; staff members from participating specialist care organizations in two group sessions based on the Promoting Action of Research Implementation in Health Services (PARIHS) framework [32, 33]; and LEAP. These consultations aimed to: gauge site readiness for recovery-oriented practice; identify potential facilitators and obstacles to implementing recovery-oriented practice in specialist care settings; examine existing organizational activities or service frameworks that could be modified to support the application of the intervention; determine supervisory strategies that would best facilitate staff uptake of the intervention and ensure the materials were inclusive of consumer issues. The adaptation process was overseen by an advisory committee of representatives from key stakeholder groups to ensure that the content and processes of the PULSAR intervention are sensitive to the Victorian mental health care system as well as the local cultural and legal contexts. Drawing on qualitative analysis of the consultation group transcripts and the advisory committee's expertise, the adaptation process addressed the following issues: 'How training is delivered' (e.g. contextualization of training, training over time, follow-up training and practical tools to keep recovery on the everyday landscape), training content related to REFOCUS elements (e.g. language, listening, common understanding of terms, building on staff strengths) and how staff could access a consumer's choices and preferences differentiating between language and processes, and a concern over the term coaching. Once these adaptations were agreed upon by the advisory committee, the training materials were considered by LEAP and their changes incorporated by the advisory committee.

The intervention focuses on promoting recovery-based practices to staff that are in addition to standard care, and is comprised of two core components: Recovery-Promoting Relationships and Working Practices.

Recovery-promoting relationships

According to a recovery-oriented framework, the working relationship between staff and consumers is crucial to the process of recovery. The intervention develops and supports this relationship by: assisting teams to develop a shared understanding of personal recovery; exploring existing values held by individual workers and the team; developing skills in coaching; and raising the expectations held by consumers that their values, strengths and goals will be prioritised in their relationships with staff members.

Working practices

The intervention is centered around three main working practices that form the specific behaviours and recovery supports necessary for building positive, recovery-promoting relationships in mental health care: 1) Understanding values, treatment and support preferences; 2) Assessing and working with strengths; and 3) Supporting goal-striving. Staff are trained to ensure that care planning is based on the consumer's values, preferences, strengths, and personally valued goals.

The intervention is supported by four implementation strategies: 1) Personal recovery training; 2) Coaching and working practice training; 3) Team manager reflection group; and 4) Team reflection sessions, as well as a set of training materials and compatible working tools. The intervention content and implementation strategies are described in detail in the PULSAR training manual, which is available from the corresponding author upon request.

After receiving the intervention, staff are invited to take part in monthly hour long PULSAR Active Learning Sessions (PALS) with an experienced PULSAR facilitator to discuss and reflect upon their experiences of delivering recovery-oriented practice in the service setting. The sessions support the practice-based implementation of the intervention through providing an

Table 3 Schedule of enrolment, interventions, and assessments

	Time points		
Project events	T0	T1	T2
Specialist staff enrolment			
Eligibility screen	X		
Informed consent	X		
Randomization	X		
Intervention			
Year 1 clusters		X	
Year 2 clusters			X
PALS			
Year 1 clusters		X	X
Year 2 clusters			X
Consumer recruitment			
Eligibility screen	X	X	X
Survey packs to eligible consumers	X	X	X
Informed consent	X	X	X
Consumer (quantitative) assessment			
cRCT - stream 1			
Demographics	X	X	X
QPR	X	X	X
cRCT – streams 2 and 3	X	X	X
WEMWBS	X	X	X
INSPIRE	X	X	X
PNCQ	X	X	X
GAF	X	X	X
SOFAS	X	X	X
CSQ	X	X	X
MASS	X	X	X
Coercion Ladder	X	X	X
Routinely collected data extracted from service medical files		X	X
Staff qualitative sub-study			
Informed consent		X	X
Individual interview		X	
Individual interview/focus group			X
Consumer qualitative sub-study			
Informed consent		X	X
Individual interview		X	
Individual interview/focus group			X
Process evaluation sub-study[a]			
Examination of specific quantitative & qualitative data in study			X
Source key study documentary notes			X
Examine staff training evaluation sheets		X	X

Notes. For a description of the T0, T1 and T2 time points, see Fig. 1. For an expansion of cRCT stream acronyms see Table 4

interactive and collaborative learning environment for staff, and ongoing access to PULSAR trainers and training resources.

Intervention modifications

The delivery of the intervention was modified to account for previously unknown restrictions on the ability of services to release staff for two days of training. In response to these restrictions, the first intervention round for clinical services was developed as a two-day session, with the community services training planned as a separate two-day session in the same week. In addition to the consumer trainer being employed by the project, trainers were sourced from clinical services for the clinical sessions and the community sector for the community sessions. This was anticipated as enabling the inclusion of specialist skills and experience in the delivery of training.

Training in the second round was planned to be subject to further modifications based on analyses of evaluations of the first round of training by both participants and trainers.

Intervention dosage

Staff movements are tracked at intervention sites every 3 months from the end of training in order to measure the degree of intervention received or "dosage". Forms are emailed to site managers every 3 months which requests that site managers list any changes in team members who have or have not undergone the intervention training, including changes to work hours and movements within the organization or externally to other organizations. All employed staff of the services are included as of the end of training census date, whether they were trained with PULSAR or not and whether they were on leave or not. This dosage information will be used in the study analyses.

Control

The control condition is standard treatment, which is defined as follows:

Monash Health: routine care as governed by the policies and procedures applicable to Monash Health, and which are consistent with the National Standards for Mental Health Services 2010 & Directives as issued from time to time by the Chief Psychiatrist of Victoria and concordant with the Mental Health Act 2014.

MHCSS: a non-clinical module of care which already has a number of elements concordant with recovery-oriented practice, and which we will be exploring whether can be further improved by the PULSAR intervention.

Measures

Both quantitative and qualitative data collection occurs in this study, see Tables 1 and 4. The primary and secondary outcome measures were chosen as they are consumer-rated measures of personal recovery and well-being. Since personal recovery is something experienced rather than assessed by an expert, self-report measures were appropriate for the study end-point.

Primary outcome measure

The primary outcome measure is the Questionnaire about the Process of Recovery (QPR [34]), a 22-item consumer-rated questionnaire used to assess experience of personal recovery. The QPR comprises of two subscales: Intrapersonal recovery processes (17 items) and Interpersonal recovery processes (5 items), with each item being rated on a 5-point Likert scale ranging from 0 (disagree strongly) to 4 (agree strongly). A higher score indicates increased recovery [34]. The QPR subscales have good internal consistency (Intrapersonal: $r = 0.94$; Interpersonal: $r = 0.77$), test-re-test reliability (Intrapersonal: $r = 0.874$, $p = 0.001$; Interpersonal: $r = 0.769$, $p = 0.001$), and construct validity [34].

Secondary outcome measures

There are two secondary consumer-rated outcomes. The 27-item Importance of services in recovery questionnaire (INSPIRE) assesses recovery support from a worker [35]. The two sub-scales of INSPIRE are: Supporting personally defined recovery (Support sub-scale; 20 items) and Working relationships (Relationship sub-scale; 7 items). Items in the Support sub-scale are first rated for whether they are important for the consumer's recovery (Yes/No). If rated Yes, the item is additionally rated on either a 5-point Likert scale ranging from 0 (Not at all) to 4 (Very much) or as 'I do not want support from my worker with this'. The Relationship sub-scale is rated on a 5-point Likert scale ranging from 0 (Strongly disagree) to 4 (Strongly agree). The measure is scored by converting the mean of all Likert ratings to a percentage ranging from 0 (low support) to 100 [35].

The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is a 14-item scale designed to assess functional and emotional well-being and appraise programs targeted towards the improvement of mental well-being [36]. The scale is rated on a 5-point Likert scale ranging from 1 (None of the time) to 5 (All of the time), providing a total score ranging from 14 to 70. A higher score indicates a higher level of mental well-being.

Other measures

Additional measures administered to consumers in streams 2 and 3 of the cRCT include:

Table 4 Primary, secondary and other outcome measures

Quantitative (Consumer) data collected in the cRCT	
Primary outcome	1. Questionnaire about the Process of Recovery (QPR)
Secondary outcomes	2. INSPIRE questionnaire [13] 3. Warwick-Edinburgh Mental Well-being Scale (WEMWBS) [36]
Other measures	4. Participant Demographic Record 5. Health economic record 6. Days out of role 7. Days absent from work 8. Service utilization questionnaire 9. The Perceived Need for Care Questionnaire (PNCQ) [38] 10. Client Satisfaction Questionnaire [39] 11. Mind Australia Satisfaction Survey [40] 12. The Coercion Ladder [41] 13. Global Assessment of Functioning Scale [42] 14. Social and Occupational Functioning Assessment Scale [42, 43]
15. Routinely collected information in service medical files (data in the year prior to interview):	
	Health of the Nation Outcome Scales (HoNoS; 12 item clinician-rated measure of social disability) [39];
	Basis 32 (consumer-rated);
	LSP16 (clinician-rated); and Focus of Care (clinician-rated).
	Diagnosis information
	Number of community/outpatient mental health contacts: • Care teams involved (discipline) • Location of contact • Date and time of contact • Focus of care for the above
	Number of inpatient mental health admissions: • Inpatient facility type, and Length of Stay (LOS) • Legal status e.g. involuntary admission, etc. Any other relevant mental health related data recorded in electronic file.
Qualitative (Consumers and staff) sub-study data	
Consumer qualitative data	
Individual interviews	
Focus groups	
Staff qualitative data	
Individual interviews	
Focus groups	
Process evaluation sub-study data[a]	

[a]The process sub-study assesses a specific set of study qualitative, quantitative and documentary data relating to each cluster. Includes the data collected from staff after participation in the PULSAR training.

- *Participant Demographic Record.* See Table 5 for demographic variables. Response categories for the ethnicity variable were chosen to represent the most common cultural/ethnic groups residing in the Monash Health catchment, sourced from the relevant local government websites. An abbreviated demographic record is included in stream 1 and includes sex, age, country of birth, year of arrival if born overseas, ethnicity, main language, and length of time the consumer has used mental health services at their current service site.
- *Health economic record.* This includes questions about occupation and income.
- *Days out of role.* This item assesses the impact of mental health problems on normal daily activities over the last 30 days.
- *Days absent from work.* This item captures the number of days absent from on usual work or occupation over the last 30 days due to illness or disability, and mental health problems.
- *Service utilisation questionnaire.* This includes questions about service use, including overnight stays in hospital and healthcare consultations, adapted from the 2007 Australian National Survey of Mental Health and Wellbeing [37]. Information about current prescription and non-prescription medication is also collected.
- *The Perceived Need for Care Questionnaire (PNCQ).* This measure classifies the consumers' perception of their need for care according to four levels: no need, unmet need, partially met need and met need. The PNCQ enables systematic assessment of perceptions of service delivery, especially in relation to mental health service evaluation [38].
- *Client Satisfaction Questionnaire (CSQ).* This consumer-rated measure assesses client satisfaction with the mental health services provided [39].
- *The Mind Australia Satisfaction Survey (MASS).* The MASS is a consumer-rated measure developed by Mind Australia to evaluate overall satisfaction with services provided, individual outcomes associated with service use, and the effectiveness of staff-consumer partnerships in mental health care service delivery [40].
- *The Coercion Ladder.* This visual analogue ladder scale provides a measure of consumers' perception of coercion in their mental health service interactions including both a hospital and community services version [41].
- *The Global Assessment of Functioning Scale (GAF).* The GAF is a researcher-rated measure of an individual's level of social, occupational and psychological functioning. The scale ranges from 0 to 100 with a lower score indicating a lower level of functioning [42].
- *The Social and Occupational Functioning Assessment Scale (SOFAS).* This researcher-rated measure provides an indication of an individual's level of functioning that is not directly influenced by the severity of a psychological condition and

Table 5 Individual and cluster-level variables available for multivariable analysis

Variable	Description
Individual level	
Demographics	
Sex	Sex of consumer.
Age	Age of consumer at survey completion date.
Country of birth	Country of birth of consumer.
Year of arrival	Year of arrival in Australia if born overseas
Ethnicity	Ethnic or cultural group that the consumer identifies with.
Main language	Main language spoken at home.
Marital Status	Marital status of consumer.
Children	Number and age of any children.
Living situation	Current living situation of consumer.
Education	Education level of the consumer.
Highest qualification	Highest qualification attained by the consumer.
Mental health service use	Length of time consumer has used mental health services.
Health economics	
Employment	Current working status of the consumer.
Income	Usual weekly income of consumer, after tax, from all sources of employment and all sources excluding paid work.
Days out of role	Number of days in the past month that the consumer was totally or partly unable to carry out normal activities because of mental health problems.
Days absent from work	Number of days in the past month that the consumer was absent from work due to illness or disability, and due to mental health problems.
Medication information	Prescription and non-prescription medications taken regularly by the consumer.
Hospitalizations	Number of hospital admissions for physical problems and for mental health problems, including number of nights in total and reasons for most recent admissions.
Consultations with health professionals	Number and length of consultations with health professionals for physical health and mental health problems.
Other	'Other' measures that are listed as 9 to 15 in Table 4 above may be investigated as independent variables when relevant. For example, the primary outcome recovery (QPR) scores may be explored for associations with the individual-level scores on the Cohesion Ladder.
Cluster level	
Cluster group	Allocated to receive the intervention at either Step 1 or Step 2.
Intervention status (0/1)	A lag time of 6 months is anticipated until intervention effects are possible. The intervention status variable indicates that this lag time has passed.
Dosage (%)	Intervention dosage.
Time since intervention	All data are time-stamped in relation to the time the intervention was received at the cluster. Time value of "0" is given for the plus/minus 3 months from date of training; "1" for 4-to-6 months post training; "2" for 7-to-9 months post training, etc. Time value of "-1" for 4-to-6 months before training; "-2" for 7-to-9 months before training, etc.
Time	Study month that survey was completed: "1" = month 1, "2" = month 2, etc.
Cluster types - stratification variables, see Table 2	
Crisis assessment & treatment team	
Mobile support and treatment service or Community Care Unit	
Community Mental Health Service	
Continuing Care team	
PARC residential facility (Mind Australia)	
PARC residential facility (Ermha)	
Community outreach service	

includes impairments caused by both physical and mental health conditions. The scale ranges from 0 to 100 with a lower score indicating a lower level of functioning [42, 43].

Consumers participating in stream 3 who report having no contact with their mental health service in the previous 12 months do not complete measures pertaining to service evaluation (CSQ, MASS, Coercion Ladder: community services version, INSPIRE).

Routinely collected information in service files

For participants in streams 2 and 3 of the cRCT, data will also be extracted from routinely collected medical records maintained by participating organisations. Data will be extracted for the 12 months prior to participation in stream 2 or 3. The inclusion of routinely collected data is intended to minimize the burden on participants by reducing the amount of measures that are administered in face-to-face interviews and to enable a detailed understanding of health service and medication use over time.

For Monash Health mental health consumers, this information will be obtained from the organization's Health Information Services scanned medical records, and will include: information about diagnosis and mental health status (such as ratings on any clinician measures); occasions of contact with services; occasions spent in residential facilities operated by the service; Health of the Nation Outcome Scales (HoNoS; 12 item clinician-rated measure of social disability) [44]; Basis 32 (consumer-rated); LSP16 (clinician-rated); and Focus of Care (clinician-rated), see Table 4.

Some of the above mentioned data routinely recorded in files of Monash Health consumers, for example the HoNoS, are not available in files of consumers from Mind Australia and Ermha. Therefore in these files we will extract service activity information collected from the respective clinical databases including information about diagnosis and mental health status (ratings on any clinician measures); occasions of contact with services; and occasions spent in residential facilities operated by the service.

Diagnosis information extracted from medical files will be used to identify participants who will be invited into stream 3 of the cRCT. Stream 3 participants must have a diagnosis of psychosis, e.g., schizophrenia, schizoaffective disorder or bipolar disorder, see Table 1.

Sample size

The primary analysis examines QPR data from consumers in the cRCT (stream 1) and requires a total sample size of 756 consumers from 14 clusters over 3 years (see Table 1). This will be sufficient to detect a medium effect size representing a change in QPR score by 6.34, see Table 1. Secondary analyses that examine data from a subset of stream 1 consumers who participate in stream 2 of the cRCT, requires a total sample size of 252 consumers over the study period. This will be sufficient to detect medium effect sizes in the QPR and two secondary outcome measures (WEMWBS and INSPIRE), see Table 1. Additional secondary analyses to examine longitudinal data from a subset of stream 2 consumers who participate in stream 3 of the cRCT, requires a total sample size of 88 consumers over the study period. This will be sufficient to detect medium-large effect sizes in the QPR and WEMWBS and INSPIRE, see Table 1.

Sample size calculations were based on 14 clusters; intracluster correlation coefficient (ICC) of 0.05; significance level set at 0.05; power of 0.80; and available published [34, 36, 45] and unpublished (INSPIRE) data about distribution properties. All sample size calculations indicate the minimum number of participants we aim to recruit and were done using Stata statistical software *stepped-wedge* [46] Version 11, StataCorp. 2009.

Recruitment

Specialist Care Service recruitment

Specialist care services were identified by the clinical and CMHS service partners in the study. A role for the Steering group was to enable initial identification and engagement of services, followed by meetings with chief executive officers or senior managers to discuss the PULSAR study and the possibility of involvement. No services declined to participate.

Consumer recruitment

Original recruitment protocol The initial consumer recruitment strategy required local coordinators at each study site to identify potential participants from service administrative and clinical databases using a systematic quota sampling template provided by the study statistician. This method of identifying potentially eligible participants was developed to ensure consumer confidentiality. The site coordinator was then responsible for overseeing the mailing of survey packs to eligible consumers, which contained a 10-page participant information sheet and consent form (PICF), a 2-page questionnaire comprising the QPR and a simple demographic survey (Stream 1 survey), and two color-coded reply paid envelopes. Participants were instructed to return the signed consent form and questionnaire separately in their respective color-coded reply paid envelopes. This strategy was designed to protect

participant confidentiality by ensuring that participant data was returned independently of identifying contact information. A unique matching code was printed on each of the forms to allow subsequent data linkage.

The original PICF invited participants to consent to one of four levels of involvement in the study and sign and return the form accordingly. Consent levels were as follows:

- Level 1 consent refers to a participant consenting to the inclusion of a returned Stream 1 survey into the project.
- Level 2 consent refers to a participant providing additional permission for the researchers to access and use relevant routinely collected clinical data.
- Level 3 consent refers to a participant being willing to be contacted for a maximum of two project interviews.
- Future research consent refers to a participant being willing to be contacted to participate in future research.

However, of the 713 letters mailed out using this initial strategy, only 21 letters (2.9%) were returned over the subsequent 5 weeks.

Modified recruitment protocol Due to this low response rate, the consumer recruitment protocol went through a series of adaptations to facilitate greater engagement and flexibility of recruitment strategies. The primary mode of recruitment through mail out was modified to a) allow mail outs of letters of invitation to complete and return the Stream 1 survey form to all eligible consumers of the participating services from each cluster site; b) replace the 10-page PICF in the survey pack with a simple one-page consent to be contacted for a face-to-face interview form, thus requiring implied consent only for return of the mailed questionnaire and demographic form; and c) provide a $10 shopping voucher for all returned questionnaires where contact details are provided.

A range of secondary recruitment strategies to promote consumer response to the mail outs were added and flexibly employed according to the needs of sites. Strategies include, for example, having researchers, including consumer researchers, present at sites to speak about PULSAR; the use of publicity materials such as advertisements, posters or PULSAR-branded materials; and direct contact with clinicians and consumers at participating sites. Considerable care was taken to ensure, as far as reasonably possible, that recruitment strategies were consistent across time points at participating clusters.

Allocation

Sequence generation

Clusters were randomized to receive the intervention at either step 1 or step 2, see Fig. 1. We used stratified randomization to ensure that cluster types were balanced across arms, see Table 2. The method of sequence generation was by simple randomization using an online Research Randomizer for random number generation. Seven randomization keys were created that corresponded to the seven strata. The randomization was performed offsite by an independent researcher during the third quarter of 2014. Investigators, site coordinators, participants and all others are unable to change the randomization key and intervention allocation given to a site.

Blinding and procedures to minimize bias

As the intervention involves training, specialist mental health care staff are aware of the intervention condition they are allocated to. Efforts are made to maintain the blindness of research assistants for the onsite recruitment and yearly face-to-face assessments for consumers by withholding information about the allocation of training to clusters and by rotating interviewers across interview and onsite recruitment clusters from T0 and T1. After conducting interviews in streams 2 and 3 of the cRCT, research assistants are asked to classify consumer participants into an intervention condition (PULSAR training provided at their site of service in year 1 or year 2) together with any specific reasons for their response and an estimate of their level of confidence in their judgement to assess whether blindness is preserved.

Procedures adopted to minimize other sources of bias include:

- Allocation status is recorded in a separate (linked) database from the database containing process and outcome data;
- Consumer participants are not informed if cluster staff at the service they attend have received the intervention training;
- The stepped-wedge design can reduce contamination of control clusters as staff in all sites know they will eventually receive the intervention [20, 21];
- In recruitment, considerable efforts are made to minimize possible sampling bias by ensuring that all eligible consumers are given the opportunity to participate. For example, the multiple levels of involvement (see above *Levels of Consumer Participation*) are designed to offer maximum flexibility for consumers to participate based on possible fluctuations in mental health; and
- Randomization was performed offsite by an independent statistician according to the procedures described above.

Data collection

The broad data collection periods are in Fig. 1. As indicated in Tables 1, 3 and 4 and discussed earlier, consumers are offered multiple levels of involvement and can contribute both quantitative data in the cRCT and/or qualitative data in the nested qualitative sub-study. Data collection from staff occurs in the nested qualitative sub-study. The process evaluation sub-study assesses a specific set of existing qualitative, quantitative and documentary data.

cRCT data collection

In stream 1 of the cRCT, cross-sectional data are collected from mail-outs to consumers at three time points, see Table 1. Stream 1 participants can return a completed QPR/demographic survey anonymously in a provided reply-paid envelope addressed to the researchers if they wish. However, they are invited to provide their contact details on a separate one-page "Participant Contact and Consent Form" if they would like to be mailed a $10 shopping voucher. Participants can additionally indicate on the Participant Contact and Consent Form if they are willing to volunteer for other parts of the PULSAR project, such as a face-to-face interview, by signing a "Consent to Future Contact" section.

In streams 2 and 3 of the cRCT, data are collected in structured face-to-face interviews from a subset of stream 1 consumers who consent to future contact, see Table 1 and earlier section *Levels of consumer participation*. A pool of around 12–14 casual research assistants conduct the face-to-face interviews. Prior to conducting interviews, all research assistants attend a compulsory 2 day training workshop facilitated by senior PULSAR researchers. This training is conducted annually prior to commencement of fieldwork each year to train new staff and maintain the skills of continuing research assistants. Training modules include: research interviewing skills; research interviewing from the consumer perspective; risk assessment, including consumer safety, risk management, and distress management; staff safety, including aggression and risk management; communication skills; research ethics; home visit protocols; and blindness. The first two interviews with consumer participants are supervised by a senior PULSAR researcher, and research assistants are provided with verbal and written feedback at the end of each interview.

Participants who complete a face-to-face interview are required to provide full written informed consent for both the interview and to the researchers accessing routinely collected data using a revised PICF. Study interviews take around 60–90 min. At the end of T0 and T1 interviews, participants are asked whether they would be interested in completing a follow-up interview approximately 12 months later (for stream 3). If participants are willing to be contacted regarding the follow-up interview, they are asked to provide their contact details, give an indication of whether they are likely to relocate in the coming year, and provide the contact details of any friend or family member who might be able to pass on letters from the PULSAR project should they no longer be contactable. If participants are not available for a follow-up interview after re-contacting attempts have been made by the researchers, no additional data will be collected. Given that participants are the recipients of the intervention through services provided by trained specialist mental health care staff, no protocol for discontinued consumer participants is necessary.

All data are recorded on paper forms which are securely stored at the PULSAR administrative site. Procedures to ensure accuracy of data extraction include double entry from selected hard copy forms, range checks and examination of outliers.

For participants who provide Level 2 consent initially, then later all participants who complete a face-to-face interview, routinely collected medical data is extracted from organization-specific medical records, see Table 4. All identifiers are removed from the service record data and replaced with a code, enabling re-identification for the purpose of linkage with the participant's interview data.

The privacy of all participants is safeguarded in accordance with the National Statement on Ethical Conduct in Human Research [47]. All data is stored on password protected computer systems located within the secure PULSAR administration site. The study data will be stored for a minimum of 7 years, after which time it may be destroyed. Re-identification codes are only accessible to the core research team responsible for data management. It is possible that participant data may be used in a non-identifiable format in future research.

Qualitative sub-study data collection

The nested qualitative sub-study investigates mental health staff experiences of implementing recovery-oriented practices following the PULSAR intervention and the challenges involved within Australian mental health settings; it also seeks to explore consumer views of how their recovery has been supported in services where this staff training intervention has taken place. Two semi-structured interview guides for use in staff interviews and consumer interviews were developed, informed by literature on recovery-oriented practice, consumer and service provider expertise within the PULSAR Qualitative Research Steering Group and consultations with PULSAR's LEAP. These interview guides are used to conduct face-to-face or telephone interviews with mental health staff three to 4 months following the PULSAR training, and with consumers five to 6 months

following the PULSAR training. Staff interviews occur prior to consumer interviews on the basis that staff are likely to be aware of their own efforts to implement changes to practice before these become as evident to consumers. Interviews with staff initially explore their understanding of recovery-oriented practice and experiences and challenges encountered in implementing a recovery-oriented framework at a service level. Subsequent interviews will invite participating staff to reflect on the de-identified interview themes, and on facilitators and barriers to implementing recovery-oriented practice in their service settings in an interview or focus group discussion. Similarly, initial face-to-face or telephone consumer interviews focus on their views and experiences of recovery-oriented practice in services where mental health staff have received training, with subsequent interviews inviting consumers to reflect on the de-identified interview themes and on supports for their recovery within and beyond services.

Sample size for the qualitative sub-study is determined sequentially by qualitative sampling processes to ensure diverse perspectives are sought, and to maximize the richness of data obtained, for which we anticipate at least 20–24 staff participants and 20–24 consumer participants will be recruited and interviewed from across the specialist mental health care sites. Recruitment strategies rely on staff and consumers opting into the study based on a convenience sampling approach, informed by the current profile of consumers and staff at participating sites. Efforts are made through the use of varied recruitment strategies including flyers and onsite visits by the researchers to ensure that a diverse range of participants are represented in the evaluation. Following the PULSAR intervention in year 2, the selection of sites, specialist mental health staff and consumers to participate in the qualitative sub-study will be guided by the extent and depth of data gathered in the first year (e.g., whether some service types are underrepresented; whether consumers on Community Treatment Orders or staff working with these service users have been recruited).

All qualitative data are audio-recorded (subject to participant consent) or documented in handwritten notes, then transcribed for coding and analysis. Coding will employ both inductive reasoning and an explicit theoretical lens [48]. Thus, qualitative data will be coded and analysed, using a constant comparative method, to identify thematic similarities and differences in participants' views within and across participant groups. Further, given the PULSAR intervention is informed by CHIME and the REFOCUS recovery-promoting practices (14,29), this theoretical framework will also be used for coding so as to identify how these concepts and practices are spoken about and understood by participants. All transcribed data are de-identified and along with all other PULSAR data are stored in password-protected files within the restricted access electronic files of the PULSAR site.

Process evaluation data collection

Given that the recovery-oriented practice involves facilitating a process of change, a process evaluation is crucial to offer explanatory variables that may influence the outcome measures [49]. The nested process evaluation will use quantitative and qualitative data to identify contextual and organisational factors that influence the effectiveness of the intervention.

The process evaluation will provide additional data relating to clusters, drawing on the PARIHS framework dimensions of evidence, context and facilitation [32]. These data can be examined further in regression and other analyses to estimate the extent to which dimensions of readiness for, exposure to, and engagement in the PULSAR program are associated with differentials in outcome measures. The process evaluation study design adopts the recommendations of Moore et al. [50] and acts on the advice of Bhanbhro and colleagues [51] in ensuring our approach is informed by theory and evidence. The interventions in PULSAR seek to change services' orientation to recovery-oriented practice through influencing the behaviour of clinical staff and adapting the systems in which they work. As Chen and Rossi [52] suggest, we use a theory-driven evaluation approach which is not dependent on a single outcome measure to confirm or refute the effectiveness of the intervention.

Following the guidelines provided by Moore et al. [50], the process evaluation will focus on collecting data that has the potential to surface explanatory variables in the complex path between intervention and outcomes. The theoretically grounded research questions include:

1. What is the role of contextual factors (leadership/support for innovation/readiness for change/organisational support for change/commitment to change and perceived supervisor support for recovery orientation practice) on the adoption of the training and patient outcomes?
2. How does dosage (number of people trained and still working at the facility/number of people attending PALS and still practicing) affect attitude to and uptake of recovery-oriented practice?
3. What is the role of clinical context (Primary or Secondary Care/Community or Acute) on attitudes to and adoption of the intervention and client outcomes?
4. What is the relationship between pre-existing engagement in recovery-oriented practice, on attitudes to training, evaluation of training and transfer of training?

Statistical analysis

Main analysis plan for cRCT

The primary analysis involves evaluating the PULSAR training intervention at the consumer level by examining the QPR data from consumers, see Table 1. The planned data collection schedule has three main periods called T0, T1 and T2, see Fig. 1. Baseline (T0) data collection occurs in the year prior to and 3 months after the step 1 intervention is delivered. The first 3 months after intervention delivery is a period still considered relevant for baseline data collection based on the Kirkpatrick training evaluation model [53] which considers that the embedding of practice change requires a minimum of 9 months after intervention is delivered, including 3 months for consolidation and 6 months for implementation. In the next period called step 1, (T1), data collection occurs during the following 12 months. Then in the next period called step 2, (T2), data collection occurs during the following 12 months. During both T1 and T2 periods, data collection at individual clusters occurs at a minimum of 9 months after the intervention was delivered to ensure embedding of intervention practices and principles, see Fig. 1.

Descriptive statistics will be used to summarize the characteristics of the clusters at baseline and consumer-level variables at time of data collection (see Table 5). Cluster-level variables are those used in the stratified randomization, which are seven types of organizational variations (see Table 2), plus the intervention status of the cluster and the time since (or before) the start of the intervention. The ICC will be calculated and reported.

The analysis of data in a stepped-wedge cRCT is most suitably analysed in mixed-effects models [54]. The primary analysis examines the effect of PULSAR on the primary outcome (consumer-level QPR scores) using a linear mixed-effects model state 'on an intention-to-treat basis'. The model will include intervention status and time as fixed effects and clusters and consumers as random effects. Normally step one is just to examine intervention – control group differences controlling for cluster, before including covariates. An a priori model-fitting analysis strategy will involve both univariate and multivariable models to be developed based on baseline consumer and cluster-level variables considered statistically significant ($p < 0.10$) or clinically important (e.g., age, sex), see Table 5, and included in the model as fixed. Model fit will be examined by comparing AIC values.

Secondary analyses will examine the effect of PULSAR on secondary outcomes (WEMWBS and INSPIRE) using a linear mixed-effects models to compare the intervention and control periods (pre-intervention).

Estimated intervention effects will be reported as the mean outcome difference for continuous outcomes and Odds Ratio for binary outcomes between intervention and control periods. This can be described as a meta-analysis approach as (in the case of continuous data) the mean change in each cluster will be standardized by using the variance of the outcome measure within that cluster. The estimated intervention effects will be reported with 95% Confidence Intervals and p values. Analysis will be conducted using Stata V.14, StataCorp. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP, 2015.

Sensitivity analyses

A missing data analysis will investigate any patterns of missingness. For each primary and secondary outcome component with missing data, multiple imputation using multivariate regression with factors of age, gender, time, and intervention status will produce 100 estimates. Sensitivity analyses will be performed using this multiple imputation to account for missing data and then re-running the analyses. Sensitivity analyses will also include the intervention dosage variable described earlier.

Economic evaluation

Overall, costs associated with each participant will follow well established health economic principles [53], and cover direct medical costs of illness, plus the labour market effects of illness. Direct medical costs are to be calculated for prescription and other medically recommended non-prescription medications, and hospital and health service contacts. Labour market productivity losses will be imputed using the human capital approach by multiplying reported days off work due to mental illness with an individual's estimated salary using instrumentation devised by this team for a previous health economic evaluation [54]. Using only days off work due to illness to capture labour market costs captures an important aspect of the cost of illness; however, it is noted that the estimates obtained will be conservative and the true cost will be higher than what we obtain because of other effects of illness such as higher rates of non-participation in employment, or underemployment.

Leadership structure

The PULSAR project adopts a module based advisory structure, overseen by a project steering group chaired by Principal Investigator (PI), Professor Graham Meadows. Four modules guide and monitor the implementation and evaluation of the project which are chaired by different members of the senior research team. The modules include Adaptation, Implementation, Research and Dissemination.

Based on the REFOCUS project [55, 56], and consistent with the commitment to co-design [10, 11], PULSAR is also supported by LEAP, an advisory group comprising people with either lived experience of mental illness or

with experience of caring for someone with mental illness. LEAP was established at the commencement of the PULSAR project and continues to meet during the trial. It provides consumer and family/carer perspectives on the project and ongoing feedback and advice on the trial.

Specification of safety parameters

No plans were made for a premature stopping of the trial. Apart from any possible breaches to consumer confidentiality, which are classified as moderate risk, all risks to the safety of consumer and specialist staff are classified as minimal.

Safety oversight

Comprehensive project protocols have been developed to address staff safety, the management of participant distress, suicidal ideation or intent, threat to harm others, and disclosure of previously undisclosed criminal acts. These protocols are readily accessible to all PULSAR research and administrative staff and are reviewed and updated on a continuing basis throughout the trial.

Dissemination policy

Overview

PULSAR takes a multi-tiered approach to dissemination to maximize the translation of knowledge into practice. Dissemination avenues will include: publication of a training manual and associated resources; development of online resources to disseminate project materials to interested parties locally and abroad; publication of project protocols and findings from each component of the project in peer-reviewed literature; production of a regular newsletter updating stakeholders on project progress and outcomes; presentations and national and international conferences; local distribution through partner organizations in the community mental health sector in Victoria; and direct communication of project outcomes to key policy makers.

Rights

In relation to copyright issues in dissemination of findings, PI Meadows and CI Slade have agreed to highly accessible publication to maximize dissemination. Specifically, there is no plan to commercialize outputs of this work and so put barriers in the way of use by others. It has been the practice of the multiple research teams involved in the PULSAR proposal to actively seek to make materials widely available as far as possible without cost, and to place barriers in the way of others commercializing such work. For example, the London REFOCUS team have disseminated the REFOCUS intervention in free-to-access booklets and through open access journal articles. The dissemination plan will make the findings widely and readily available along with source training materials.

Discussion

The PULSAR Specialist Care trial will examine the efficacy of a recovery-oriented practice training intervention for specialist mental health care staff using a two-step stepped-wedge cRCT design. This design is often favoured for such community-based pragmatic trials, as the intervention will eventually be delivered to all participating clusters but can be implemented in stages to manage the practical constraints associated with delivering a large-scale intervention across multiple sites [20, 21]. The challenges encountered in the trial are providing valuable insights on how to facilitate staff adherence to the training and hence the embedding of the intervention into participating services, as well as effective methods for engaging and retaining the participation of consumers. A significant contribution of the work will be the production and dissemination of a package of professional training resources to support the implementation of recovery-oriented practice into community-based mental health services. Although the PULSAR materials have been developed according to the needs of the Australian mental health system and the local social, legal and cultural contexts, we anticipate that these resources will be adaptable to other settings and jurisdictions. In line with the approach taken by our UK partner, the PULSAR materials will be made widely and readily available.

With the current emphasis in mental health policy on refocusing services towards recovery, the results of this trial, including an assessment of clinical, organizational and health economic outcomes, will contribute to the small but growing evidence-base promoting the development of recovery-oriented service frameworks. If successful, it will be the most definitive trial to date in Australia demonstrating that the concept of recovery, and interventions designed to foster recovery-oriented staff behaviour and relationships with consumers, can be operationalized and comprehensively evaluated. Findings, and other information gathered and lessons learned during the trial, will support the continued transformation of the mental health sector towards recovery, ultimately leading to improved outcomes for people with serious mental illness.

Abbreviations

ANZCTR: Australian New Zealand Clinical Trials Registry; CATT: Crisis assessment and treatment team; CCT: Continuing care team; CCU: Community care unit; CI: Chief investigator; cRCT: Cluster randomized control trial; CSQ: Client satisfaction questionnaire; GAF: Global assessment of functioning scale; HoNoS: Health of the nation outcome scales; HREC: Human Research Ethics Committee; ICC: Intraclass correlation coefficient; LEAP: Lived experience advisory panel; MASS: Mind Australia Satisfaction Survey; MHCSS: Mental health community support services; MST: Mobile support and treatment team; PALS: PULSAR active learning sessions; PARC: Prevention and recovery in community; PARIHS: Promoting action of research implementation in health services; PI: Principal investigator; PICF: Participant information and consent form; PNCQ: Perceived need for care questionnaire; PULSAR: Principles unite local services assisting recovery; QPR: Questionnaire about the process of recovery; SOFAS: Social

and occupational functioning assessment scale; SPIRIT: Standard protocol items: recommendations for interventional trials; WEMWBS: Warwick-Edinburgh mental well-being scale

Acknowledgements
We acknowledge Mr. Umit Agis, A/Prof Jakqui Barnfield, Dr. Joanne Brooker, Ms. Kimberly Buck, Prof David Clarke, Dr. Margaret Grigg, A/Prof Iraklis Harry Minas, A/Prof Brendan Murphy, Dr. Mehmet Özmen and Ms. Christine Thornton. We also acknowledge Dr. Mary Leamy, Dr. Victoria Bird and other members of the REFOCUS team as well as the work of the PULAR trainers, Nita Mahbubani, Paul Carr, Carl Mast, Michelle Evers, Leonie Davey, Zoe Smith, Lisa Reed, Richard Kasperczyk and Michelle Kehoe.
This project was undertaken through Monash University.

Funding
This research was supported by the Victorian Government through the Victorian Mental Illness Research Fund.

Availability of data and materials
Only investigators and approved researchers added by ethics approval will have access to the final trial data set.

Authors' contributions
All authors participated in the design of the trial and intervention. FS provides overall coordination for field work and staff training and led the writing of the paper. LB chairs the research module task-group and provided oversight to development of all elements of the design. PW chairs the implementation module which oversaw the delivery of the training intervention. AB had a key role in project implementation and developing early drafts of the paper. JE and GM led development of key elements of the design and analysis approach. BI provides overall support and oversight for the health economic analysis. JJ, VE, LB, GM, PW and EWE developed the specific specialist care training intervention and associated resources. GM, LB, FS, VE and EWE developed the specialist care instrumentation and fieldwork trial protocols. MS, as head of the UK REFOCUS team and a PULSAR investigator, has provided critical consultation and resources supporting the design of the intervention and advice on aspects of the research methods. Key contributions regarding qualitative research, process evaluation and cross-cultural design elements were made by EF, EWE and RK respectively. All authors read, amended and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Research ethics approval
The project has approval from Monash Health (project approval number 14102B and 14427B) and Monash University (project approval number CF14/1600–2,014,000,773 and CF15/266–2,015,000,120) Human Research Ethics Committees. Approvals were granted (quantitative study: May 2014 and qualitative study: Jan 2015) before projects commenced and extend to December 2017.
Research audit
The Human Research Ethics Committee at Monash Health conducted an independent audit of PULSAR research procedures in the third year of the project, focusing on the correct usage and storage of participant data. No irregularities or deviations from the protocol were reported.
Consumer consent to participate
As described above, consumer participants are provided with the option of various levels of involvement in the project. For participants in Stream 1 of the cRCT, consent is implied by the return of the mailed questionnaire. It is presumed that those who do not wish to participate will not return the completed measures. Participants who are also willing to be contacted for a face-to-face interview (Streams 2 and 3 of the cRCT) indicate their consent by signing a consent to be contacted form and providing their contact details on the same form. Participants who complete a face-to-face interview are required to provide full written informed consent for both the interview and to the researchers accessing routinely collected data using a revised PICF. The project acknowledges that consenting to participate in research is an ongoing process. Project protocols ensure that participants in Streams 2 and 3 are provided with an overview of what their participation involves on each occasion, and are provided with the opportunity to reconfirm their consent at the time of interview/s.
In the nested qualitative sub-study, a participant information sheet is provided to interested potential participants and informed written consent is obtained from all consumer participants prior to interview. If, at any time over the duration of the project, a participant wishes to no longer participate and/or is deemed by the interviewer to no longer fully understand participation requirements, the research will cease.
Specialist mental health care staff consent to participate
Cluster staff who participate in the qualitative sub-study are provided with a participant information sheet and required to provide written informed consent prior to interview and/or focus group.
Discontinuing participants
All participants are advised on the consent form/s and by research assistants during face-to-face interviews that they can withdraw from the study at any time. Unless the discontinuing participant requests that their data is removed, data already collected will be included in analysis.

Author details
[1]Southern Synergy, Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Dandenong Hospital, 126 - 128 Cleeland St, Dandenong, Victoria 3175, Australia. [2]Royal District Nursing Service Institute, 31 Alma Rd, St Kilda, Victoria, Australia. [3]Mind Australia, Heidelberg, VIC, Australia. [4]Melbourne School of Population and Global Health, University of Melbourne, Parkville, Victoria 3010, Australia. [5]Department of Occupational Therapy, School of Primary and Allied Health Care, Monash University, Peninsula Campus, Frankston, VIC, Australia. [6]Department of Econometrics and Business Statistics, Faculty of Business and Economics, Monash University, Melbourne, Australia. [7]Graduate School of Business and Law, RMIT University, Melbourne, VIC 3001, Australia. [8]College of Business, Victoria University, Melbourne, VIC, Australia. [9]Institute of Mental Health, School of Health Sciences, University of Nottingham, Triumph Road, Nottingham NG7 2TU, UK.

Received: 13 April 2017 Accepted: 21 April 2017
Published online: 08 May 2017

References

1. Shepherd G, Boardman J, Slade M. Making recovery a reality. London: Sainsbury Centre for Mental Health; 2008.
2. Anthony WA. A recovery-oriented service system: setting some system level standards. Psychiatr Rehabil J. 2000;24(2):159–68.
3. Slade M. Personal recovery and mental illness: a guide for mental health professionals (values-based medicine). Cambridge: Cambridge University Press; 2009.
4. Glover H. Recovery, life long learning, social inclusion and empowerment: is a new paradigm emerging? In: Ryan P, Ramon S, Greacen T, editors. Empowerment, lifelong learning and recovery in mental health: towards a new paradigm. London: Palgrave Macmillan; 2012. p. 15–35.
5. Slade M. 100 ways to support recovery. 2nd ed. London: Rethink Mental Illness; 2013.
6. Anthony W. Toward a vision of recovery: for mental health and psychiatric rehabilitation services. Boston: Boston University Press; 2007.
7. Ash D, Bland R, Brown P, Browne MO, Edwards R, Emmerson B, et al. Mental health services in the Australian states and territories. In: Meadows G, Farhall JF, Fossey EM, Grigg M, McDermott FM, Singh B, editors. Mental health in Australia: collaborative community practice. 3rd ed. South Melbourne, Victoria: Oxford University Press; 2012. p. 118–54.
8. Victorian Department of Health. Framework for recovery-oriented practice. Melbourne: State Government of Victoria; 2011.
9. Australian Health Ministers. National Mental Health Plan 2003–2008. Canberra: Australian Government; 2003.
10. Co-design Initiative. Co-design: Shared Perspectives on Authentic Co-design. 2016. https://auspwn.files.wordpress.com/2016/05/codesign-shared-perspectives-report-vf1-5-040616.pdf. Accessed 14 Jul 2016.
11. Victorian Government. Victoria's 10-Year Mental Health Plan. 2015. https://www2.health.vic.gov.au/about/publications/policiesandguidelines/victorias-10-year-mental-health-plan. Accessed 25 Jan 2017.

12. Minas H, Kakuma R, Too LS, Vayani H, Orapeleng S, Prasadllides R, et al. Mental health research and evaluation in multicultural Australia: developing a culture of inclusion. Int J Ment Health Syst. 2013;7(1):23.
13. Slade M, Bird V, Le Boutillier C, Williams J, McCrone P, Leamy M. REFOCUS Trial: protocol for a cluster randomised controlled trial of a pro-recovery intervention within community based mental health teams. BMC Psychiatry. 2011;11(1):185–98.
14. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry. 2011;199(6):445–52.
15. Slade M, Bird V, Clarke E, Le Boutillier C, McCrone P, Macpherson R, et al. Supporting recovery in patients with psychosis through care by community-based adult mental health teams (REFOCUS): a multisite, cluster, randomised, controlled trial. Lancet Psychiatry. 2015;2(6):503–14.
16. Mueser KT. Supporting recovery from psychosis: assessment of REFOCUS. Lancet Psychiatry. 2015;2(6):478–9.
17. Shawyer F, Farhall J, Thomas N, Hayes SC, Gallop R, Copolov D, et al. Acceptance and commitment therapy for psychosis: randomised controlled trial. Br J Psychiatry. 2016; doi:10.1192/bjpbp116182865.
18. Enticott JC, Shawyer F, Brophy L, Russell G, Fossey E, Inder B, et al. The PULSAR Primary Care protocol: a stepped-wedge cluster randomized controlled trial to test a training intervention for general practitioners in recovery-oriented practice to optimize personal recovery in adult patients. BMC Psychiatry. 2016;16:451.
19. Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346:e7586.
20. Mdege ND, Man M-S, Taylor CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. J Clin Epidemiol. 2011;64(9):936–48.
21. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol. 2006;6(1):54–63.
22. Creswell J. A concise introduction to mixed methods research. Thousand Oaks: Sage Publications; 2014.
23. Creswell J, Clark V. Designing and conducting mixed methods research. Thousand Oaks: Sage Publications; 2007.
24. De Allegri M, Pokhrel S, Becher H, Dong H, Mansmann U, Kouyaté B, et al. Step-wedge cluster-randomised community-based trials: an application to the study of the impact of community health insurance. Health Res Policy Syst. 2008;6(1):10.
25. Edan V, Meadows G, Brophy L, Weller P, Thornton C. PULSAR research: making mental health services more recovery-orientated. New Paradigm: Aust J Psychosoc Rehabil. 2015;46:46–8.
26. Health M. Fact Facts 2013–2014. Melbourne: Monash Health; 2014.
27. City of Greater Dandenong. A profile of health and wellbeing in Greater Dandenong. Victoria: Greater Dandenong Council; 2013.
28. Browne G, Hemsley M. Consumer participation in mental health in Australia: what progress is being made? Australas Psychiatry. 2008;16(6):446–9.
29. Jorm AF, Kelly CM, Morgan AJ. Participant distress in psychiatric research: a systematic review. Psychol Med. 2007;37(7):917–26.
30. Bird V, Leamy M, Le Boutillier C, Williams J, Slade M. REFOCUS: Promoting recovery in community mental health services. London: Rethink; 2011.
31. Slade M, Bird V, Le Boutillier C, Farkas M, Grey B, Larsen J, et al. Development of the REFOCUS intervention to increase mental health team support for personal recovery. Br J Psychiatry. 2015;207(6):544–50.
32. Rycroft-Malone J. The PARIHS framework - a framework for guiding the implementation of evidence-based practice. J Nurs Care Qual. 2004;19(4):297–304.
33. Stetler CB, Damschroder LJ, Helfrich CD, Hagedorn HJ. A guide for applying a revised version of the PARIHS framework for implementation. Implement Sci. 2011;6(1):99.
34. Neil ST, Kilbride M, Pitt L, Nothard S, Welford M, Sellwood W, et al. The questionnaire about the process of recovery (QPR): a measurement tool developed in collaboration with service users. Psychosis. 2009;1(2):145–55.
35. Williams J, Leamy M, Bird V, Le Boutillier C, Norton S, Pesola F, et al. Development and evaluation of a measure to identify mental health service support for recovery (INSPIRE). Soc Psychiatry Psychiatr Epidemiol. 2015;50:777–86.
36. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes. 2007;5:63.
37. Australian Bureau of Statistics. National Survey of Mental Health and Wellbeing: summary of results (4326.0). Canberra: Australian Bureau of Statistics; 2008.
38. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). Soc Psychiatry Psychiatr Epidemiol. 2000;35(9):427–35.
39. Attkisson CC, Zwick R. The Client Satisfaction Questionnaire: psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann. 1982;5(3):233–7.
40. Wray M, Brophy L, Loh M. The Prevention and Recovery Care Exit Survey. Mind Australia: Victoria; 2013.
41. Høyer G, Kjellin L, Engberg M, Kaltiala-Heino R, Nilstun T, Sigurjónsdóttir M, et al. Paternalism and autonomy: a presentation of a Nordic study on the use of coercion in the mental health care system. Int J Law Psychiatry. 2002;25(2):93–108.
42. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington: American Psychiatric Association; 2000.
43. Goldman H, Skodol A, Lave T. Revising Axis V for DSM-IV: a review of measures of social functioning. Am J Psychiatry. 1992;149(9):1148–56.
44. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and Development. Br J Psychiatry. 1998;172(1):11–8.
45. Lindsay G, Davies H, Band S, Cullen MA, Cullen S, Strand S, et al. Parenting Early Intervention Pathfinder Evaluation. UK: University of Warwick; 2009.
46. Hemming K, Marsh J. A menu-driven facility for sample-size calculations in cluster randomized controlled trials. Stata J. 2013;13(1):114–35.
47. The National Health and Medical Research Council, The Australian Research Council, The Australian Vice-Chancellors' Committee. National Statement on Ethical Conduct in Human Research. Canberra: Commonwealth of Australia; 2007. (Updated May 2015)
48. Thorne S. Data analysis in qualitative research. EBN Notebook. 2000;3:68–70.
49. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions an introduction to the new Medical Research Council guidance. Evidence–Based Public Health: Effectiveness and Efficiency. 2010:185–203.
50. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ. 2015;350:1–7.
51. Bhanbhro S, Gee M, Cook S, Marston L, Lean M, Killaspy H. Recovery-based staff training intervention within mental health rehabilitation units: a two-stage analysis using realistic evaluation principles and framework approach. BMC Psychiatry. 2016;16(1):292.
52. Chen H, Rossi P. Evaluating with sense: the theory-driven approach. Eval Rev. 1983;7:283–302.
53. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs. The Four Levels. 3rd ed. San Franciso: Berrett-Koehler Publishers; 2006.
54. Barker D, McElduff P, D'Este C, Campbell MJ. Stepped wedge cluster randomised trials: a review of the statistical methodology used and available. BMC Med Res Methodol. 2016;16:69.
55. Slade M, Bird V, Chandler R, Fox J, Larsen J, Tew J, et al. The contribution of advisory committees and public involvement to large studies: case study. BMC Health Serv Res. 2010;10(1):323.
56. Slade M, Slade M, Trivedi P, Trivedi P, Chandler R, Chandler R, et al. Developing involvement during a programme of recovery research. JMHTEP. 2016;11(4):244–55.