



OPEN ACCESS

EDITED BY

Zahiruddin Quazi Syed,
Datta Meghe Institute of Medical Sciences,
India

REVIEWED BY

Tomi Bergström,
University of Jyväskylä,
Finland
Dorothea Jäckel,
Vivantes Hospital,
Germany

*CORRESPONDENCE

Stephen Pilling
✉ s.pilling@ucl.ac.uk

SPECIALTY SECTION

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

RECEIVED 21 October 2022

ACCEPTED 03 February 2023

PUBLISHED 22 February 2023

CITATION

Alvarez-Monjaras M, Lotmore M, Razzaque R,
Hopfenbeck MS and Pilling S (2023) The
community mental health team fidelity scale: A
measure of program fidelity of social networks
interventions for severe mental illness.
Front. Psychol. 14:1076791.
doi: 10.3389/fpsyg.2023.1076791

COPYRIGHT

© 2023 Alvarez-Monjaras, Lotmore, Razzaque,
Hopfenbeck and Pilling. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

The community mental health team fidelity scale: A measure of program fidelity of social networks interventions for severe mental illness

Mauricio Alvarez-Monjaras¹, Melissa Lotmore¹,
Russell Razzaque², Mark Steven Hopfenbeck³ and
Stephen Pilling^{1*}

¹Department of Clinical Educational and Health Psychology, University College London, London, United Kingdom, ²North East London NHS Foundation Trust, London, United Kingdom, ³Department of Health Sciences, Norwegian University of Science and Technology, Trondheim, Sør-Trøndelag, Norway

Open dialogue (OD) is a multi-component therapeutic and organizational intervention for crisis and continuing community mental health care with a therapeutic focus on clients' social networks. The development and implementation of this model of care in the United Kingdom requires considerable contextual adaptations which need to be assessed to support effective implementation. Program fidelity—the extent to which core components of an intervention are delivered as intended by an intervention protocol at all levels—is crucial for these adaptations.

Aims: To develop, pilot, and implement a program fidelity measure for community mental health services providing OD and 'treatment as usual' (TAU) or standard NHS crisis and community care.

Methods: Measure structure, content, and scoring were developed and refined through an iterative process of discussion between the research team and OD experts. Measure was piloted in the 6 OD and 6 TAU services participating in a large-scale research program.

Results: Initial data suggests that the Community Mental Health Team Fidelity Scale (COM-FIDE) is a potentially reliable and feasible measure of the fidelity of community mental health services and specific OD components of such services.

KEYWORDS

open dialogue, fidelity, implementation science, community mental health, measure development, severe mental illness, complex interventions

Introduction

Poor social networks have been associated with both the development and maintenance of mental illness (Giacco et al., 2012). Interventions targeting social networks—such as the Open Dialogue (OD) approach (Seikkula et al., 1995) might therefore help ameliorate mental health crises and reduce the likelihood of relapse. However, due to limited staff training and skills, and a lack of continuity associated with the current model of crisis and continued community care of the British National Health Service (NHS), such interventions are not currently provided (Razzaque and Wood, 2015; The Commission on Acute Adult Psychiatric Care, 2015). Further, the professional and contextual adaptations required to integrate OD successfully and sustainably

into NHS models of care require a consideration of the model's core components.

Program fidelity or the extent to which core components of an intervention are delivered as intended by a treatment protocol is a useful approach to supporting effective implementation (Santacrose et al., 2004; Borelli, 2011; Gearing et al., 2011). This paper outlines the development, piloting, and implementation of a program fidelity measure for the OD approach: The Community Mental Health Team Fidelity Scale (COM-FIDE). The paper begins with a brief description of Open Dialogue and the current NHS model of crisis and continuing community care in mental health. This is followed by an exploration of some of the challenges involved in integrating OD into the provision of mental health services in the United Kingdom, including the challenges in developing a fidelity measure. The COM-FIDE development and piloting method are then outlined, alongside some preliminary psychometric data. Finally, results are considered alongside the utility of COM-FIDE.

Crisis and continuing community mental health care in the United Kingdom

The NHS is facing significant problems in providing care and support for people with severe mental illness, potentially due to poorly developed and increasingly fragmented pathways of care (NHS Confederation, 2016; The Kings Fund, 2016). This is in part a consequence of the functional model of mental health care, where care is often provided by several different teams, each with its own criteria for acceptance (Morton and Norman-Nott, 2019). Standard NHS crisis and continuing community care services for people experiencing severe mental illness consist primarily of crisis resolution and home treatment teams (CRTs) and community mental health teams (CMHTs). As an alternative to hospitalization, these multidisciplinary teams—typically conformed by psychiatrists, mental health nurses, social workers, and support workers—provide intensive assessment, care, and support in patients' homes (Weisman, 1989; Jethwa et al., 2007; Johnson, 2013). Standard care often acknowledges and may attempt to work with the social network of a person in crisis; however, their brief and functional nature and the pressures on service resources make this form of ongoing network-oriented care a challenging endeavor (Razzaque and Wood, 2015).

Despite the promise shown in randomized controlled trials (Johnson et al., 2005a,b; Lloyd-Evans et al., 2014, 2019), questions have been raised on whether standard care might be decreasing in effectiveness (Johnson et al., 2005a,b; Jacobs and Barrenho, 2011). Wheeler et al. (2015) suggested this might be due to a considerable atrophy of its key functions, with many services offering limited home visits outside of office hours and only 50% of services providing post-hospital discharge care. It is important to ask whether this possible decrease in the quality of community-based services can be explained by a lack of resources or if organizational problems, such as staff competencies, roles, care pathways, or fidelity to a model, may also be contributory factors.

Open Dialogue

Open Dialogue (Seikkula et al., 1995) is both a therapeutic approach and a way of organizing mental health services developed in

Finland, which explicitly targets social networks. The aim of Open Dialogue is to promote a greater shared understanding of service users' problems, a greater sense of agency, collaborative decision-making, and the network's mutual support in the long term (Seikkula et al., 1995, 2006; Seikkula et al., 2001a, 2011). This is done through the enactment of the principles of (1) immediate help, (2) social networks perspective, (3) flexibility and mobility, (4) responsibility, (5) psychological continuity, (6) tolerance of uncertainty, and (7) dialogue and polyphony (Seikkula et al., 1995). In contrast to current models of care—in which families may not be directly involved—Open Dialogue uses network meetings attended by family members, friends, and other professionals involved with the service user as the central means of intervention delivery (Seikkula et al., 1995; Seikkula and Olson, 2003; Lakeman, 2014; Razzaque and Wood, 2015). Service users and their social networks engage in shared decision-making with healthcare professionals to agree on appropriate pharmaceutical, psychological, or social interventions (Seikkula et al., 2006; Olson et al., 2014).

The development of an integrated OD approach to the provision of mental health services offers the possibility of an alternative to the current 'functional team' model of care in the United Kingdom (Hopfenbeck, 2015; Razzaque and Stockmann, 2016). Preliminary evidence suggests that OD may be more effective than standard care in reducing relapse and the use of antipsychotic medication (Seikkula et al., 2001b, 2003; Hartman and De Coursey, 2015; Bergström et al., 2018). Additionally, OD might help equip mental health staff with additional skills necessary to engage service users and their families across the broad spectrum of care needs (Holmesland et al., 2014). However, although promising, there is no high-quality evidence to date to support an NHS-wide adoption of this model of care.

Program fidelity measurement

Transferring Open Dialogue from one health care setting to another requires considerable contextual adaptations that could undermine structural (i.e., organizational) and process (i.e., therapeutic) components of the original model (Gonzalez Castro et al., 2004). In fact, international OD implementation programs (e.g., Pocobello and Salamina, 2015) have noted that the organizational change is such, that staying faithful to the OD principles (e.g., Seikkula et al., 2006; Olson et al., 2014) has encountered significant obstacles. Program fidelity or the extent to which an intervention is delivered as intended in a treatment protocol at all levels can be a useful tool for understanding an intervention's critical components on a structural, organizational, and functional level (Carroll et al., 2007; Proctor et al., 2011; Teague et al., 2012).

Literature suggests that program fidelity measures should involve (1) an evidence-based, comprehensive, and multimodal approach to assessment, (2) clearly and objectively operationalized components stemming from a coherent and comprehensive theory of change, and (3) easily-available data from the relevant stakeholders (Schoenwald et al., 2011; Essock et al., 2015). Although uncommon, existing measures for multi-component interventions such as OD are somewhat consistent in terms of measure design, assessment procedures, and scoring. Donabedian (1988) suggested a structure-process-outcome framework for fidelity evaluation; however, most measures emphasize structural features of service provision (e.g., operations, staffing, or services provided) but tend to neglect

important process and outcome features relevant to the therapeutic model (Alvarez-Monjaras, 2019).

A few efforts have been made to establish appropriate fidelity measures for standard crisis and continuing community care. The CORE CRT (Lloyd-Evans et al., 2016) is the most robust and validated measure to date for crisis services. However, OD implementation studies so far have focused on practitioner adherence or the quality of delivery of network meetings according to the key OD principles (Eiterå et al., 2014; Olson et al., 2014; Rambøll, 2014; Ziedonis et al., 2018; Lotmore et al., 2022). Since OD is not only a therapeutic model but also a way of organizing care, it is important to identify not only the clinically relevant (i.e., process) features but also the structural and organizational features that characterize the approach and distinguish it from standard care. In other words, if OD is to be successfully implemented and integrated into the traditional NHS model of crisis and continued community care, it is essential to develop a program fidelity measure to support the implementation of OD that is faithful not only to the original Finnish model, but also fit for its incorporation into the NHS.

Study aims

This study was part of the NIHR ODDESSI (Open Dialogue: Development and Evaluation of a Social Network Intervention for Severe Mental Illness) program grant (RP-PG-0615-20,021). ODDESSI aims to evaluate whether OD –when integrated within standard NHS mental health services for adults in crisis–improves the clinical and cost-effectiveness of standard crisis and continuing community mental health care (i.e., CRTs and CMHTs). The ODDESSI is a cluster-randomized controlled trial (RCT) consisting of five work packages oriented toward defining, implementing, and evaluating OD services across 28 trial clusters from five NHS trusts (for full protocol see Pilling et al., 2022).

The key goal of the present study was to develop, pilot, and implement a program fidelity measure that could accurately characterize the quality of both standard NHS crisis and continuing community care (hereafter referred to as ‘treatment as usual’ or TAU) and high-quality OD practice. If successful, this measure would provide information on whether: (1) NHS services, once reorganized on an OD model of care, can deliver OD with sufficient fidelity to its core principles and ensure they are both provided effectively; (2) it is possible to distinguish OD services from standard care based on their model of work; and (3) there are any differences in implementation between each model’s teams.

Methods

Study design

Although this specific study was relevant to all work packages of the ODDESSI trial, it was embedded in the second work package as part of the feasibility stage (WP2). WP2 addressed the feasibility of a cluster RCT, including the question of whether adherence and fidelity measures could provide a reliable measure of OD practice. Additionally, the NIHR shared their concern that—in order to draw meaningful conclusions from the outcomes—the trial needed to be able

to compare OD teams against high-quality TAU. Developing a program fidelity measure is one way of ensuring high quality of care in both OD and TAU.

Participants

Participants for this study were staff members from six OD services and their six TAU counterparts. For each service, one pair of managers and one pair of practitioners (i.e., psychiatrists, psychologists, psychotherapists, nurses, social workers, and support workers) were interviewed by two joint independent raters. A total of 48 staff members were interviewed.

Ethical approval

This study received ethics approval from the Health Research Authority under reference number 18/LO/0026. No personal or confidential information was solicited. Participants gave consent for being recorded using an encrypted and password-protected recorder.

Measure development

The Community Mental Health Team Fidelity Scale (COM-FIDE) was developed following a stepwise approach (Bond et al., 2000; Holmbeck and Devine, 2009), based on our systematic review of existing measures (Alvarez-Monjaras, 2019), and a series of discussions with experts (Figure 1).

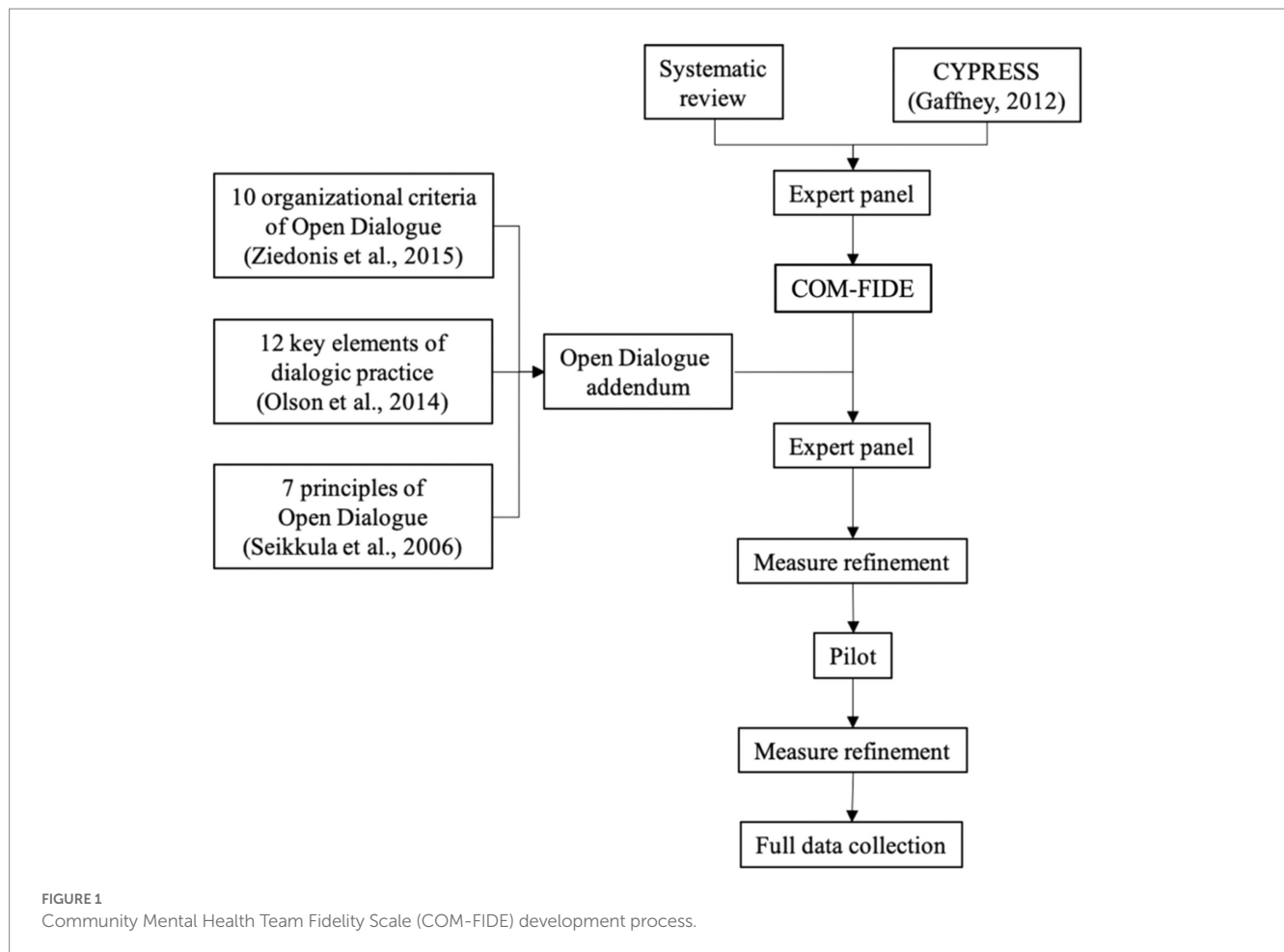
Defining the content and scope of the measure

The initial content, method of delivery, and scoring process of the COM-FIDE builds on work done at University College London on the ‘Children and Young People – Resource, Evaluation and Systems Schedule’ (CYPRESS) (Gaffney, 2012) and findings from our systematic review. CYPRESS was developed for the Systemic Therapy for At-Risk Teens (START) RCT (Fonagy et al., 2013) to characterize services delivering multisystemic therapy and management as usual for young people with complex presentations. CYPRESS captures key elements of effective implementation efforts (e.g., coherent theoretical basis, high program fidelity, qualified staff, sustained approach, etc.) across three levels of service delivery: service characteristics, team operations, and delivery of interventions. The promising results from the START trial suggested that CYPRESS could be a robust measure for service characterization.

Drawing on the CYPRESS (Gaffney, 2012), our systematic review, and Donabedian’s (1988) structure-process-outcome framework, the research team agreed to four broad key domains to assess: (1) team structure and culture, (2) access to and engagement with services, (3) delivery of care, and (4) external support. An initial list of items was drafted and then refined based on three factors: (1) a focus on adult mental health, (2) the ability to encompass both OD and TAU, and (3) the ability to identify high-quality TAU.

Designing the measure

The refinement and detail of the measure outline was established through a series of meetings and discussions between the main author



(MA, Clinical Psychologist), SP (Clinical Psychologist with expertise in evidence-based practice and experience in measure development), RR (Consultant Psychiatrist, ODDESSI co-applicant, and international expert in OD), MH (Lead OD trainer), and ML (Clinical Psychologist involved in the development of the adherence measure). An iterative process, aimed at achieving an acceptable level of utility of the measure, took place between October 2017 and January 2018.

Open Dialogue fidelity

Another important goal of WP2 was to refine the OD protocol for its implementation across NHS sites. Consequently, the resulting measure needed to be able to recognize features specific to open dialogue in OD teams. A similar item development process for an OD addendum took place based on existing OD literature (e.g., Seikkula et al., 2006; Olson et al., 2014; Ziedonis et al., 2015), and the ODDESSI treatment protocol. The ODDESSI protocol set out key functions, referral pathways, and governance arrangements of each site, and was developed by the research team in collaboration with experts in OD and TAU, alongside senior NHS staff and clinicians.

Given the complexity of OD terminology, a series of discussions around the main theoretical principles (e.g., dialogism, transparency, openness) were arranged with OD experts. The aim was to determine the best possible way to translate these key principles into objective and reliable service-level items that could be ascertained by raters not trained in OD.

The community mental health team fidelity scale

The above led to the development of a 25-item Community Mental Health Team Fidelity Scale (COM-FIDE, formerly CoMFideS). COM-FIDE is a measure designed to describe the structure, functioning, pathways, community links, and delivery of care provided by good quality community MH services, including OD. COM-FIDE is hence a measure of program fidelity of both (a) standard NHS crisis and continued community care and (b) best practice in OD delivery. The COM-FIDE also includes a 7-item Open Dialogue Addendum focused on measuring the level of fidelity to open dialogue principles of care.

The first section of the COM-FIDE concerns structural aspects of the services under assessment. The COM-FIDE comprises four sections that assess the level of fidelity of mental health teams—regardless of their model of care—to high-quality crisis and continued community care: (1) Team structure and culture (8 items); (2) Access and engagement (6 items); (3) Delivery of care (6 items); and (4) External and support (5 items).

Measure refinement

The COM-FIDE was piloted in one OD and one TAU service to identify areas of improvement in the COM-FIDE and assess the

measure's acceptability. For each team, two managers and two practitioners were interviewed using the draft measure, followed by a brief discussion on its structure, content, and acceptability. Using the outcome from the pilot, the measure was once again refined and discussed with the expert panel (See [Figure 1](#)).

The COM-FIDE manual

Based on initial work gathered throughout the measure development process and drawing from CYPRESS ([Gaffney, 2012](#)), a manual was developed (Available on the UCL ODDESSI website). The manual includes a description and rationale for each item alongside their respective behavioral anchors for scoring ([Alvarez-Monjaras and Pilling, 2019](#)).

Measure pilot

Recruitment and setting up the interviews

As per the manual, initial contact with services was done *via* email correspondence, describing the study, its purpose, and a brief description of the measure and interview process. Service documentation (e.g., staffing, supervision, safeguarding, and operational policies) was also requested from each Trust to gather service-level data. Interviews typically lasted no more than 60 min. The average time spent per interview was 46 min (range = 35–57). None of the raters were OD-trained but were all clinicians trained on the use of COM-FIDE, which included discussions of each item and rating examples.

Agreeing on final ratings

Once each interview session was completed, both raters reviewed their individual score sheets separately. Each item was then jointly reviewed to identify and record disagreements and a consensus reached on the final score.

Scoring and cut-off scores

During each interview, both raters simultaneously and independently rate a copy of the COM-FIDE. Once finished, they review and reach a consensus in the ratings. All items of the COM-FIDE are rated on a 4-point behaviorally anchored Likert scale. Advised that a 4-tier structure might offer the highest level of precision possible for rating program fidelity, rather than the traditional 5-point Likert approach. On all items, a score of one indicates that the principle at hand is not present or there is insufficient evidence of its enactment in the team's way of functioning, whereas a score of 4 indicates that the principle is enacted or carried out in an excellent manner and with no visible shortcomings or inconsistencies across the team. The overall COM-FIDE score yields to a final score of 100 and the OD-addendum to a score of 28. Each section then obtains an average score of its composite items (for more information on scoring and all behavioral anchors, please refer to the manual on the UCL ODDESSI website).

Providing (1) this is the first fidelity measure developed for open dialogue in the NHS, and (2) that there are no pre-existing criteria for what constitutes a 'good' standard of TAU care, nor of open dialogue fidelity, we considered 4 fidelity gradations: an average score on each section equal or above 3.40 (85th percentile) was considered 'very good'; scores between 2.80 and 3.39 (70–85th percentiles) as 'good';

scores between 2.40 and 2.79 (60–69th percentiles) as 'acceptable'; and scores equal or below 2.39 (below 60th percentile) as 'poor' or lacking fidelity.

Data analysis

Data from each site consisted of: (1) three rating sheets (i.e., two independent rating sheets and a final rating sheet) for manager interviews, and (2) three rating sheets for practitioner interviews. Data from all rating sheets were entered into an Excel spreadsheet and later exported onto an SPSS database. All analyzes were conducted using IBM SPSS Statistics version 25 for Mac ([IBM Corporation, 2017](#)). Descriptive statistics and radar plots were used at service level to characterize site fidelity scores. Statistical tests comparing scores were not conducted given the small sample size.

Psychometric properties

The present study explored—albeit tentatively—the following psychometric properties of the COM-FIDE: (1) inter-rater reliability, (2) internal consistency, and (3) face and content validity.

Reliability

Reliability analyzes were based on item-level data from the independent rating sheets. In terms of inter-rater reliability, Pearson's *r* or intra-class coefficients (ICC) were not obtained given the sample size, and that respondents and raters were not fully crossed or nested. Neither of these tests can remove systematic coder deviations and can therefore underestimate the true reliability of ill-structured measurement designs (ISMDs) such as the one used for this study ([Putka et al., 2008](#); [Hallgren, 2012](#)). The G estimation coefficient ([Putka et al., 2008](#)) was chosen to make up for the limited data and as a less biased reliability estimator. A G coefficient above 0.7 was considered acceptable. Internal consistency reliability was assessed using Cronbach's alpha ([Cronbach, 1951](#)). Alpha coefficients above 0.7 were considered acceptable ([Streiner, 2003](#)).

Validity

Face and content validity were assumed as adequate given the iterative feedback and input from experts, managers, and staff members. Other forms of measure validity were not considered given the scarcity of data.

Results

Service characteristics

All TAU and OD interviews were completed with no missing data. Only TAU teams were able to provide copies of their operational policies as OD teams were still in the process of developing their own; however, given the structure of the trial clusters, TAU policies were also taken to apply to OD teams. The average caseload per staff member was 25.8 service users (SD = 7.36, range = 20–40) for the OD teams and 29.8 (SD = 8.50, range = 25–45) for TAU teams. The mean staff for OD teams was 9.5 (SD = 3.08, range = 5–13) and for TAU teams was 13.8 (SD = 3.49, range = 10–19). Psychiatrists, psychiatric nurses, clinical psychologists, and psychotherapists were the most common professions and were all

TABLE 1 Service characteristics.

	Open Dialogue (n=6)		Standard care (n=6)	
	\bar{X} (Range)		\bar{X} (Range)	
Employed staff (FTE and WTE)	9.50 (5–13)		13.82 (10–19)	
Caseload	\bar{X} (SD)		\bar{X} (SD)	
Team	220.83 (120.68)		503.33 (165.73)	
Individual	25.83 (7.36)		29.83 (8.50)	
	<i>n</i>	%	<i>n</i>	%
Service setup				
Integrated	5	83.3	0	0.0
Stand-alone	1	16.7	6	100.0
Staff roles				
Psychiatrists	6	100.0	6	100.0
Nurses	6	100.0	6	100.0
Nurse assistants	2	33.3	1	16.7
Psychologists	6	100.0	6	100.0
Occupational therapists	3	50.0	5	83.3
Social workers	3	50.0	4	66.7
Support workers	3	50.0	5	83.3
Peer support workers	6	100.0	1	16.7
Advocates/volunteers	0	0	1	16.7
Weekly team meetings	6	100.0	6	100.0
Supervision arrangements				
Individual	5	83.0	6	100.0
Group	6	100.0	3	50.0

employed across teams (n=6). Occupational therapists were employed by 83% (n=5) of TAU teams, whereas only in 50% of OD teams. Only one TAU team (8%) employed advocates. Nurse assistants were employed by 25% of the teams (n=3) altogether (Table 1).

Preliminary psychometric properties of the COM-FIDE measure

Reliability analysis

Item-level calculations of the G estimate of reliability suggested a potentially good inter-rater reliability across the measure. All but two items showed coefficients above 0.6, and 17 of the 32 items (53.1%) showed coefficients above 0.9 (Table 2). The item ‘Flexibility of Response’ had a reliability coefficient of 0.42 and the item ‘OD continued professional development’ had a coefficient of 0, given its null variance (rate variance=0.000, rater variance=0.000, estimated variance of the combination of rate*rater interaction and residual effects=2.298).

Internal consistency

Both the 25-item COM-FIDE scale and the 7-item OD addendum suggested potentially good internal consistency, with Cronbach’s alpha coefficients of 0.90 on the overall COM-FIDE scale and 0.95 in the OD addendum (see Table 3 for subscale-specific coefficients). An item-level analysis was conducted to examine whether deleting any

individual item would make important changes to the overall internal consistency of each scale. Results suggested little influence of any individual item on the total internal consistency of the 25-item COM-FIDE scale (coefficient change ranging from -0.002 to 0.01) and the 7-item OD addendum (range = -0.020–0.016).

When analyzed on a section level, all 5 sections appeared to have adequate internal consistency (Table 3). Results suggested little influence of any individual item on the total internal consistency of their respective section (coefficient increases ranging from 0.02 to 0.04 across sections); however, some items showed very small item-total correlations (minimum value of 0.3; Field, 2017). Further, some items were found to negatively correlate with their sub-samples. For instance, in the ‘Team structure and culture’ section, items ‘Supervision’ had a negative item-total correlation of -0.01 as did ‘Training’ and ‘Staff roles’, with coefficients of -0.25 and -0.29, respectively. Also, in the ‘Access and engagement’ section, item ‘Flexibility of response’ had a negative item-total correlation of -0.04. All other items had item-total correlation coefficients above 0.4.

Item scores

On an item level, 6 of the 25 COM-FIDE items (24%) had mean scores equal to or above 3.40 (‘very good’); 14 items (56%) had scores between 2.80 and 3.39 (‘good’); two items (8%) had scores between 2.40 and 2.79 (‘acceptable’); and three items (12%) had scores below 2.39 (‘poor’; Table 4).

TABLE 2 Inter-rater reliability of the COM-FIDE using the G estimate (n=24).

Item	G(0.200, 2)
COM-FIDE scale	0.992
Team structure and culture	
1. Team ethos and comprehensiveness	0.914
2. Staff training	0.868
3. Supervision	0.829
4. Staff roles	0.918
5. Team capacity	0.897
6. Routine outcome monitoring	0.952
7. Safety	0.896
8. Service-user involvement in co-production	0.944
Access and engagement	
1. Access to the service	0.927
2. Providing information	0.689
3. Prompt action	0.818
4. Identification of support systems	0.916
5. Flexibility of response	0.421
6. Assertive engagement	0.913
Delivery of care	
1. Continuity of care	0.896
2. Establishing clinical meetings	0.918
3. Collaborative decision making	0.950
4. Information sharing and communication	0.751
5. Service-user involvement in the delivery of care	0.829
6. Coordination of care	0.646
External support	
1. Service linkage	0.884
2. Community links (Practitioner level)	0.783
3. Community links (Support system)	0.929
4. Caregiver involvement and support	0.969
5. Discharge and aftercare	0.760
Open dialogue addendum	
1. Transparency	0.929
2. Self-disclosure	0.970
3. Intervention frequency	0.990
4. Intervention content and structure	0.995
5. Team self-work	0.964
6. OD training	0.995
7. OD continued professional development	0.000

Standard of care (COM-FIDE score)

Overall, the mean COM-FIDE total score (i.e., excluding the OD addendum, as this dimension was only relevant to OD teams) across all 12 teams was 3.11 (SD=0.38, range=2.72–3.72), possibly suggesting ‘good’ fidelity to standard NHS care. When analyzed by

TABLE 3 Internal consistency of COM-FIDE subscales.

COM-FIDE subscale (n=24)	Internal consistency (Cronbach’s alpha)
Team structure and culture	0.681
Access and engagement	0.677
Delivery of care	0.817
External support	0.713
Open Dialogue addendum	0.954

model of care, the 6 OD teams had a mean COM-FIDE total score of 3.25 (SD=0.38; range=2.78–3.72), whereas the 6 TAU teams had a mean COM-FIDE total score of 2.97 (SD=0.35, range=2.72–3.66). Open dialogue teams had higher scores in all sections compared to TAU teams (Figure 2).

Overall, OD teams scored higher on most items. TAU teams scored higher than OD teams in ‘co-production’ (mean=2.25, SD=0.52), ‘service capacity’ (mean=2.92, SD=0.49) ‘routine outcome measurement’ (mean=2.17, SD=0.26), ‘access to the service’ (mean=3.08, SD=0.66), and ‘prompt action’ (mean=3.58, SD=0.58; Figure 3).

Open Dialogue fidelity

When focusing only on the 6 OD teams, three of the 6 teams (50%) showed ‘very good’ fidelity, 2 teams (33%) were in the ‘good’ range, and one team (17%) demonstrated ‘acceptable’ fidelity. On an item level, 4 of the 7 items (57.1%) had mean scores equal to or above 3.40 (‘very good’); two items (14.2%) had scores between 2.80 and 3.39 (‘good’); and one item (14.2%) had scores between 2.40 and 2.79 (‘acceptable’; Figure 4).

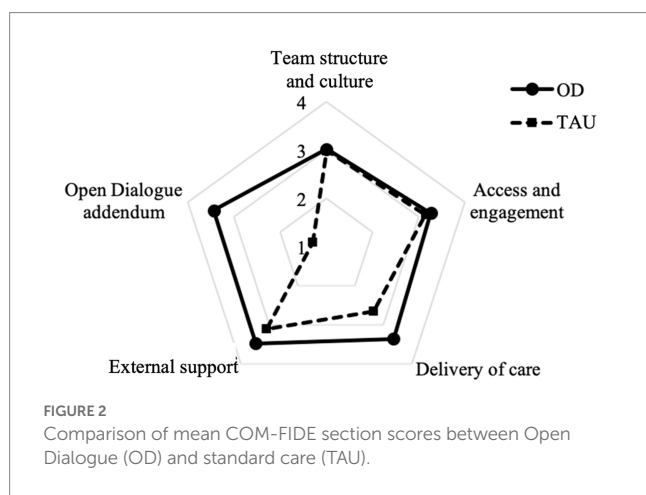
Discussion

The Community Mental Health Team Fidelity Scale

These preliminary findings suggest that COM-FIDE is a robust measure of program fidelity for crisis and continued community care teams aiming at integrating OD to their practice. This is in line with the findings on the CYPRESS measure (Gaffney, 2012), which was shown to be a robust measure for assessing MST fidelity. As noted by Waters et al. (2021) in a recent discussion paper, there are significant commonalities between COM-FIDE general scale and the CYPRESS scale. Both were developed by the same research group and designed to capture all core components of well-functioning community-based teams (in the case of CYPRESS, for services for children and young people). The design of COM-FIDE supports its use as a measure of fidelity for current standard community care (the comparator in many evaluations). In establishing the ODDESSI program, the research group drew a distinction between OD as an organizational intervention (i.e., fidelity) which is measured by the COM-FIDE measure, and a therapeutic intervention (i.e., adherence) which is measured by the OD Adherence Scale (Lotmore et al., 2022). While this approach will require additional reviews we believe this is more

TABLE 4 Differences in COM-FIDE mean scores between service models (n=12).

	Open Dialogue (n=6)		Standard care (n=6)	
	Mean (SD)	Range	Mean (SD)	Range
COM-FIDE score	3.25 (0.38)	2.78–3.72	2.97 (0.35)	2.72–3.66
Team structure and culture	3.02 (0.37)	2.56–3.44	2.99 (0.35)	2.63–3.63
Access and engagement	3.26 (0.40)	2.58–3.75	3.15 (0.44)	2.58–3.83
Delivery of care	3.35 (0.51)	2.67–4.00	2.65 (0.48)	2.17–3.50
External support	3.47 (0.34)	3.10–3.90	3.10 (0.44)	2.60–3.70
Open dialogue addendum	3.44 (0.36)	2.93–3.79	1.30 (0.30)	1.00–1.86



than compensated for by allowing for the key organizational elements of OD and standard care to be robustly compared. The adherence measure is only of relevance for OD services.

In terms of reliability, inter-rater reliability is promising especially considering that none of the raters were OD-trained. Although three items of the general scale and one item from the OD addendum were below acceptable ranges, it is possible that this was a consequence of unclear behavioral anchors. Interestingly, both ‘providing information’ and ‘coordination of care’ received mixed feedback from experts. Developers argued that providing information about the service to clients and referrers helps streamline access to the service; however, there were some doubts on whether these two features were too similar to tease apart during interviews. Similarly, coordination of care was considered a key component of crisis and continued community care; however, there were concerns about this item being redundant. With regards to ‘flexibility of response’, the low reliability may have been due to the lack of clarity in the definition, which made it difficult for raters to reach a consensus in scores. As per the lack of inter-rater variance in the OD-specific item of “continued professional development,” this may have been because all OD sites attended the same CPD programme and anchors were not sensitive enough to identify major differences in extended training beyond percentages of staff engagement. Future versions of the manual could include clearer definitions and more specific behavioral anchors.

In terms of validity, COM-FIDE appears to have adequate content validity and the ODDDESSI team considered it feasible for use in the full trial. The iterative item refinement process, as well as the discussions with international experts in the field (including the

developer of Open Dialogue) were central to developing items that would fit both models of care while also being sufficiently sensitive to possibly distinguish between them.

Defining a ‘good’ standard of care

Results suggest that all teams demonstrated a ‘good’ standard of care against the criteria set out above. Most OD teams developed from TAU teams (except for one team which was an independent team prior to the trial); with a varying degree of experience, staffing, and capacity across teams.

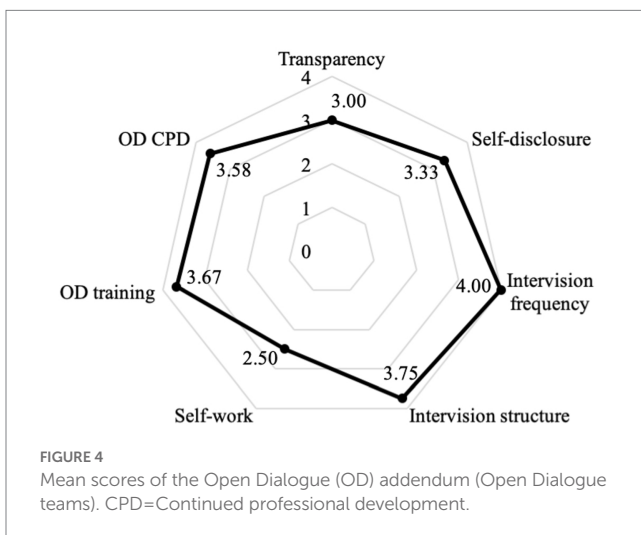
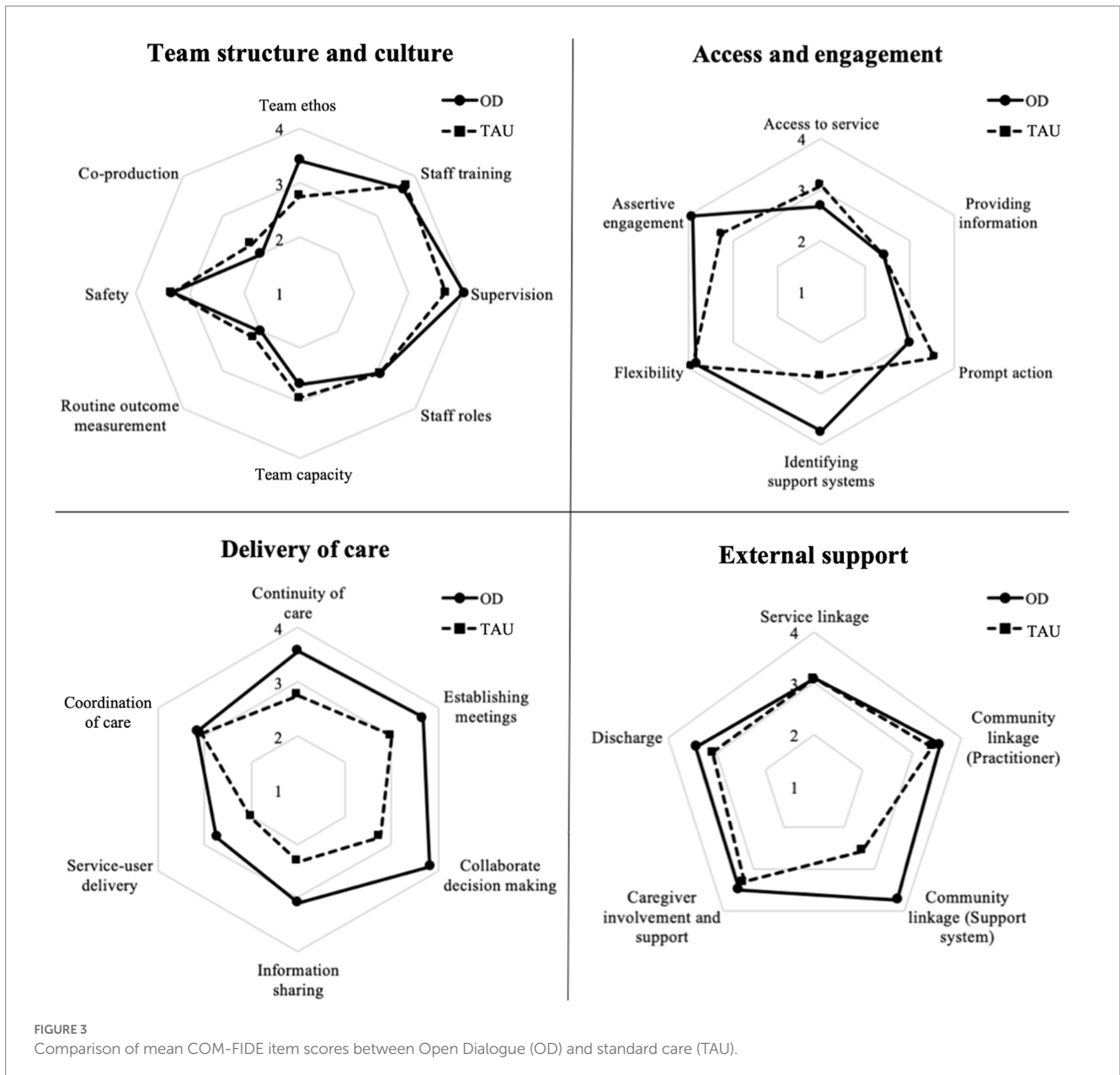
The four-tier cut-off approach was chosen as a solution based on existing literature on fidelity measurement. Although it proved useful in determining whether participating teams were ready for inclusion in the trial (i.e., ‘acceptable’ fidelity) it was limited in setting variations in fidelity above the cut-off. Further data collected across a range of OD and a range of community mental health teams should support further refinement of the scale.

Strengths and limitations

COM-FIDE is a feasible and reliable measure for use in the ODDDESSI program and is the first measure to explicitly address service level delivery of open dialogue. Its development and results from the present study identified a number of strengths but also highlighted some limitations of the measure.

One of the main strengths of the study is in the measure development process. One of the aims of the ODDDESSI trial was to comprehensively assess the organizational and therapeutic elements of OD by developing valid and reliable measures to compare OD versus current standard care. We believe this was best achieved through two distinct measures (i.e., COM-FIDE and OD Adherence Scale). Having the opportunity to discuss and revise the measure with experts in the field allowed for a rich discussion about the theoretical ‘critical components’ of the Finnish OD approach to translate the therapeutic principles (Seikkula et al., 1995), and key elements (Olson et al., 2014; Ziedonis et al., 2015) into measurable structural and therapeutic variables. A modified Delphi approach to expert feedback (Dalkey and Helmer, 1963) may have nonetheless provided more structure to the measure development process.

In terms of limitations, a larger sample would have allowed for more robust methods (e.g., factor analysis); but as noted



above this could be addressed when additional data becomes available. Another limitation was that raters were not fully crossed or nested given the difficulties in matching respondent and rater availability. This limitation was addressed in two ways: first, the G estimator—although unconventional—seemed a robust solution to this as it considers rater assortment and systematic rater deviations; and as interviews were recorded it is possible to further assess reliability using novel independent raters.

Conclusion

This paper describes the development, piloting, and testing of a program fidelity measure for its use in the ODESSI program. The Community Mental Health Team Fidelity Scale (COM-FIDE) captures both standard NHS crisis care practice and open dialogue practice. The measure development process used recognized

methods including multiple raters, multiple data sources, and multiple settings to assess its properties. Preliminary psychometric results were encouraging, suggesting that COM-FIDE is suitable for use in a range of community mental health settings. Results suggest that COM-FIDE may be able to establish (a) the extent to which teams deliver their respective models according to their protocols, and (b) the degree of differentiation between similar approaches to crisis care and recovery.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by this study received ethics approval from the Health Research Authority under reference number 18/LO/0026. No personal or confidential information was solicited. Participants gave consent for being recorded using an encrypted and password-protected recorder. The patients/participants provided their written informed consent to participate in this study.

Author contributions

MA-M, ML, and SP contributed to conception and design of the study. MA-M organized the database, performed the statistical analysis, and wrote the first draft of the manuscript. All authors contributed to data collection, manuscript revision, read, and approved the submitted version.

References

- Alvarez-Monjaras, M. (2019). *Fidelity Measurement for the Implementation of Social Networks Interventions in Complex Mental Health*. (Doctoral Dissertation, University College London).
- Alvarez-Monjaras, M., and Pilling, S. (2019). *The Community Mental Health Fidelity Scale Rating Manual*. Unpublished Manuscript. Available at: <https://www.ucl.ac.uk/pals/research/clinical-educational-and-health-psychology/research-groups/oddesi/related-references>
- Bergström, T., Seikkula, J., Alakare, B., Mäki, P., Kögäs-Saviaro, P., Taskila, J., et al. (2018). The family-oriented open dialogue approach in the treatment of first-episode psychosis: nineteen-year outcomes. *Psychiatry Res.* 270, 168–175. doi: 10.1016/j.psychres.2018.09.039
- Bond, G., Williams, J., Evans, L., Salyers, M., Kim, H., Sharpe, H., et al. (2000). *Psychiatric Rehabilitation Fidelity Toolkit*. Cambridge, MA: Human Services Research Institute.
- Borelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J. Public Health Dent.* 71, S52–S63. doi: 10.1111/j.1752-7325.2011.00233.x
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., and Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement. Sci.* 2:40. doi: 10.1186/1748-5908-2-40
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: 10.1007/BF02310555
- Dalkey, N., and Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Manag. Sci.* 9, 458–467. doi: 10.1287/mnsc.9.3.458
- Donabedian, A. (1988). The quality of care: how can it be assessed? *J. Am. Med. Assoc.* 260, 1743–1748. doi: 10.1001/jama.1988.03410120089033
- Eitèr, A., Hansen, L., Vind, B., Hansen, A., Castella, J., and Sørensen, E. (2014). *Open Dialogue—Key Markers and their Context*. [Unpublished].
- Essock, S., Nossel, I., McNamara, K., Bennet, M., Buchanan, R., Kreyenbuhl, J., et al. (2015). Practical monitoring of treatment fidelity: examples from a team-based intervention for people with early psychosis. *Psychiatr. Serv.* 66, 674–676. doi: 10.1176/appi.ps.201400531
- Field, A. (2017). *Discovering Statistics Using SPSS. 5th*. London: SAGE.
- Fonagy, P., Butler, S., Goodyer, I., Cottrell, D., Scott, S., Pilling, S., et al. (2013). Evaluation of multisystemic therapy pilot services in the systemic therapy for at risk teens (START) trial: study protocol for a randomised controlled trial. *Trials* 14, 265–219. doi: 10.1186/1745-6215-14-265
- Gaffney, C. (2012). *Complex Interventions for Children and Young People: Exploring Service Delivery Frameworks and Characterising Interventions*. (Doctoral Dissertation, University College London). Available at: <http://discovery.ucl.ac.uk/1367068/1/C.GAFFNEY%20VOLUME%201.pdf>
- Gearing, R., El-Bassel, N., Ghesquiere, A., Baldwyn, S., Gillies, J., and Ngeow, E. (2011). Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin. Psychol. Rev.* 31, 79–88. doi: 10.1016/j.cpr.2010.09.007
- Giacco, D., McCabe, R., Kallert, T., Hansson, L., Fiorillo, A., and Priebe, S. (2012). Friends and symptom dimensions in patients with psychosis: a pooled analysis. *PLoS One* 7:e50119. doi: 10.1371/journal.pone.0050119
- Gonzalez Castro, F., Barrera, M., and Martinez, C. (2004). The cultural adaptation of prevention interventions: resolving tensions between Fidelity and fit. *Prev. Sci.* 5, 41–45. doi: 10.1023/B:PREV.0000013980.12412.cd
- Hallgren, K. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor. Quant. Methods Psychol.* 8, 23–34. doi: 10.20982/tqmp.08.1.p023

Funding

This work was supported by the National Institute for Health Research (NIHR) under grant RP-PG-0615-20021, the NIHR UCLH Biomedical Research Center, and the Mexican National Council for Science and Technology (CONACYT) under scholarship 600356. The views presented are those of the authors and may not reflect those of the funding agencies who did not have input into the manuscript.

Acknowledgments

Parts of this work have been submitted in partial fulfillment of the requirements for the Doctorate in Clinical Psychology, at University College London and exist in the online repository of theses of the institution.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer TB declared a past co-authorship with the author MA-M to the handling editor.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Hartman, D., and De Coursey, J. (2015). Family therapy in the real world: dialogical practice in a regional Australian public mental health service. *Aust. N. Z. J. Fam. Ther.* 36, 88–101. doi: 10.1002/anzf.1088
- Holmbeck, G., and Devine, K. (2009). Editorial: an author's checklist for measure development and validation manuscripts. *J. Pediatr. Psychol.* 34, 691–696. doi: 10.1093/jpepsy/jsp046
- Holmesland, A., Seikkula, J., and Hopfenbeck, M. (2014). Inter-agency work in open dialogue: the significance of listening and authenticity. *J. Interprof. Care* 28, 433–439. doi: 10.3109/13561820.2014.901939
- Hopfenbeck, M. (2015). Peer-supported open dialogue. *Context* 138, 29–31.
- IBM Corporation. (2017). *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM.
- Jacobs, R., and Barrenho, E. (2011). Impact of crisis resolution and home treatment teams on psychiatric admissions in England. *Br. J. Psychiatry* 199, 71–76. doi: 10.1192/bjp.bp.110.079830
- Jethwa, K., Galappathie, N., and Hewson, P. (2007). Effects of a crisis resolution and home treatment team on in-patient admissions. *Psychiatr. Bull.* 31, 170–172. doi: 10.1192/pb.bp.106.010389
- Johnson, S. (2013). Crisis resolution and home treatment teams: an evolving model. *Adv. Psychiatr. Treat.* 19, 115–123. doi: 10.1192/apt.bp.107.004192
- Johnson, S., Nolan, F., Hoult, J., White, I., Bebbington, P., Sandor, A., et al. (2005a). Outcomes of crises before and after introduction of a crisis resolution team. *Br. J. Psychiatry* 187, 68–75. doi: 10.1192/bjp.187.1.68
- Johnson, S., Nolan, F., Pilling, S., Sandor, J., Hoult, J., McKenzie, N., et al. (2005b). Randomised controlled trial of acute mental health care by a crisis resolution team: the North Islington crisis study. *BMJ* 331:599. doi: 10.1136/bmj.38519.678148.8F
- Lakeman, R. (2014). The Finnish open dialogue approach to crisis intervention in psychosis: a review. *Psychother. Aust.* 20, 28–35.
- Lloyd-Evans, B., Bond, G., Ruud, T., Ivanekka, A., Gray, R., Osborn, D., et al. (2016). Development of a measure of model fidelity for mental health crisis resolution teams. *BMC Psychiatry* 16:427. doi: 10.1186/s12888-016-1139-4
- Lloyd-Evans, B., Christoforou, M., Osborn, D., Ambler, G., Marston, L., Lamb, D., et al. (2019). *Crisis resolution teams for people experiencing mental health crises: the CORE mixed-methods research programme including two RCTs*. Southampton, UK: Programme Grants for Applied Research, 1–102.
- Lloyd-Evans, B., Mayo-Wilson, E., Harrison, B., Istead, H., Brown, E., Pilling, S., et al. (2014). A systematic review and meta-analysis of randomised controlled trials of peer support for people with severe mental illness. *BMC Psychiatry* 14, 14–39. doi: 10.1186/1471-244X-14-39
- Lotmore, M., Ziedonis, D., Alvarez-Monjaras, M., Hopfenbeck, M., Razzaque, R., Wilson, E., et al. (2022). Development and refinement of the open dialogue (OD) adherence protocol in complex mental health care. *Front. Psychol.* 13:1041375. doi: 10.3389/fpsyg.2022.1041375
- Morton, C., and Norman-Nott, A. (2019). "The interface with community services" in *The Oxford Textbook of Inpatient Psychiatry*. eds. A. Barrera, C. Attard and R. Chaplin (England: Oxford University Press).
- NHS Confederation. (2016). *Is Mental Health Crisis in Crisis?* (Issue Brief 290). Available at: <https://www.nhsconfed.org/media/confederation/files/publications/documents/is-mental-health-crisis-care-in-crisis-mhn-briefing-290.pdf>.
- Olson, M., Seikkula, J., and Ziedonis, D. (2014). *The Key Elements of Dialogic Practice in Open Dialogue: Fidelity Criteria*. The University of Massachusetts Medical School. Worcester, MA.
- Pilling, S., Clarke, K., Parker, G., James, K., Landau, S., Weaver, T., et al. (2022). Open dialogue compared to treatment as usual for adults experiencing a mental health crisis: protocol for the ODESSI multi-site cluster randomised controlled trial. *Contemp. Clin. Trials* 113:106664. doi: 10.1016/j.cct.2021.106664
- Pocobello, R., and Salamina, G. (2015) Introducing Open Dialogue in Italy: The Role of Participatory Evaluation. In «Listen To Me!» *Humanizing Human Practices-Third International Conference on Dialogical Practices*. Norway: Kristiansand. Available at: <https://goo.gl/NAVp4e>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm. Policy Ment. Health Serv. Res.* 38, 65–76. doi: 10.1007/s10488-010-0319-7
- Putka, D., Le, H., McCloy, R., and Diaz, T. (2008). Ill-structured measurement designs in organizational research: implications for estimating interrater reliability. *J. Appl. Psychol.* 93, 959–981. doi: 10.1037/0021-9010.93.5.959
- Rambøll. (2014). *ABEN dialog Del II-Manual*. Aarhus: Rambøll.
- Razzaque, R., and Stockmann, T. (2016). An introduction to peer-supported open dialogue in mental healthcare. *BJPsych Adv.* 22, 348–356. doi: 10.1192/apt.bp.115.015230
- Razzaque, R., and Wood, L. (2015). Open dialogue and its relevance to the NHS: opinions of NHS staff and service users. *Community Ment. Health J.* 51, 931–938. doi: 10.1007/s10597-015-9849-5
- Santacroce, S., Maccarelli, L., and Grey, M. (2004). Intervention fidelity. *Nurs. Res.* 53, 63–66. doi: 10.1097/00006199-200401000-00010
- Schoenwald, S., Garland, A., Chapman, J., Frazier, S., Sheidow, A., and Southam-Gerow, M. (2011). Toward the effective and efficient measurement of implementation fidelity. *Adm. Policy Ment. Health* 38, 32–43. doi: 10.1007/s10488-010-0321-0
- Seikkula, J., Aaltonen, A., Alakare, B., Haarakangas, K., Keränen, J., and Lehtinen, K. (2006). Five-year experience of first-episode nonaffective psychosis in open-dialogue approach: treatment principles, follow-up outcomes, and two case studies. *Psychother. Res.* 16, 214–228. doi: 10.1080/10503300500268490
- Seikkula, J., Aaltonen, A., Alakare, B., Haarakangas, K., Keränen, J., and Sutela, M. (1995). "Treating psychosis by means of open dialogue" in *The Reflecting Team in Action: Collaborative Practice in Family Therapy*. ed. S. Friedman (New York: The Guilford Press)
- Seikkula, J., Alakare, B., and Aaltonen, A. (2001a). Open dialogue in psychosis I: an introduction and case illustration. *J. Constr. Psychol.* 14, 247–265. doi: 10.1080/107205301750433397
- Seikkula, J., Alakare, B., and Aaltonen, A. (2001b). Open dialogue in psychosis II: a comparison of good and poor outcome cases. *J. Constr. Psychol.* 14, 267–284. doi: 10.1080/107205301750433405
- Seikkula, J., Alakare, B., and Aaltonen, A. (2011). The comprehensive open-dialogue approach in Western Lapland: II. Long-term stability of acute psychosis outcomes in advanced community care. *Psychosis* 3, 192–204. doi: 10.1080/17522439.2011.595819
- Seikkula, J., Alakare, B., Aaltonen, J., Holma, J., Rasinkangas, A., and Lehtinen, V. (2003). Open dialogue approach: treatment principles and preliminary results of a two-year follow-up on first episode schizophrenia. *Ethical Hum. Sci. Serv.* 5, 163–182.
- Seikkula, J., and Arnkil, T. (2006). *Dialogical Meetings in Social Networks*. London: Karnak.
- Seikkula, J., and Olson, M. (2003). The open dialogue approach to acute psychosis: its poetics and micropolitics. *Fam. Process* 42, 403–418. doi: 10.1111/j.1545-5300.2003.00403.x
- Streiner, D. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J. Pers. Assess.* 80, 99–103. doi: 10.1207/S15327752JPA8001_18
- Teague, G., Mueser, K., and Rapp, C. (2012). Advances in fidelity measurement for mental health services research. *Psychiatr. Serv.* 63, 765–771. doi: 10.1176/appi.ps.201100430
- The Commission on Acute Adult Psychiatric Care. (2015) *Old Problems, New Solutions: Improving Acute Psychiatric Care for Adults in England*. Royal College of Psychiatrists: England, London.
- The Kings Fund. (2016). *Deficits in the NHS 2016 (Briefing)*. Available at: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Deficits_in_the_NHS_Kings_Fund_July_2016_1.pdf
- Waters, E., Ong, B., Mikes-Liu, K., McCloughen, A., Rosen, A., Mayers, S., et al. (2021). Open dialogue, need-adapted mental health care, and implementation fidelity: a discussion paper. *Int. J. Ment. Health Nurs.* 30, 811–816. doi: 10.1111/inm.12866
- Weisman, G. (1989). "Crisis intervention" in *A Clinical Guide for the Treatment of Schizophrenia*. ed. A. Bellack (Boston, MA: Springer).
- Wheeler, C., Lloyd-Evans, B., Churchard, A., Fitzgerald, C., Fullarton, K., Mosse, L., et al. (2015). Implementation of the crisis resolution team model in adult mental health settings: a systematic review. *BMC Psychiatry* 15, 74–14. doi: 10.1186/s12888-015-0441-x
- Ziedonis, D., Olson, M., and Seikkula, J. (2015). *10 Organizational Criteria for Open Dialogue*. The University of Massachusetts Medical School. Worcester, MA.
- Ziedonis, D., Small, E., and Larkin, C. (2018). *Dialogic Practice Fidelity Rating Manual*. San Diego, California: University of California, San Diego.