ORIGINAL PAPER

Shared decision-making in mental health care using routine outcome monitoring: results of a cluster randomised-controlled trial

Margot J. Metz[1,2,3] · Marjolein A. Veerbeek[2] · Jos W. R. Twisk[4] · Christina M. van der Feltz-Cornelis[1,5] · Edwin de Beurs[6,7] · Aartjan T. F. Beekman[4,8]

Abstract

Purpose To investigate the effects of Shared Decision-Making (SDM) using Routine Outcome Monitoring (ROM) primary on patients' perception of Decisional Conflict (DC), which measures patients' engagement in and satisfaction with clinical decisions, and secondary on working alliance and treatment outcomes.

Method Multi-centre two-arm matched-paired cluster randomised-controlled trial in Dutch specialist mental health care. SDM using ROM (SDMR) was compared with Decision-Making As Usual (DMAU). Outcomes were measured at baseline (T0) and 6 months (T1). Multilevel regression and intention-to-treat analyses were used. Post hoc analyses were performed on influence of subgroups and application of SDMR on DC.

Results Seven teams were randomised to each arm. T0 was completed by 186 patients (51% intervention; 49% control) and T1 by 158 patients (51% intervention, 49% control). DC, working alliance, and treatment outcomes reported by patients did not differ significantly between two arms. Post hoc analyses revealed that SDMR led to less DC among depressed patients ($p = 0.047$, $d = -0.69$). If SDMR was applied well, patients reported less DC (SDM: $p = 0.000$, $d = -0.45$; ROM: $p = 0.021$, $d = -0.32$), which was associated with better treatment outcomes.

Conclusion Except for patients with mood disorders, we found no difference between the arms for patient-reported DC. This might be explained by the less than optimal uptake of this generic intervention, which did not support patients directly. Regarding the positive influence of a higher level of applying SDM and ROM on less DC and better treatment outcomes, the results are encouraging for further investments in patient-oriented development and implementation of SDMR.

Keywords Shared decision-making · Routine outcome monitoring · Outcome measurements · Decisional conflict · Cluster randomised-controlled trial

Introduction

Several studies have shown beneficial effects of Shared Decision-Making (SDM) in mental health care [1–5]. Patients report that they are better informed and more actively engaged in treatment, experience higher satisfaction, and reduced decisional conflict, and patients show more treatment adherence. This can lead to better treatment outcomes. SDM is the collaborative approach in which patients and clinicians share available information

✉ Margot J. Metz
m.metz@ggzbreburg.nl

1 GGz Breburg, Postbus 770, 5000 AT Tilburg, The Netherlands

2 Trimbos Institute, Postbus 725, 3500 AS Utrecht, The Netherlands

3 VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

4 VU University Medical Centre Amsterdam, Postbus 7057, 1007 MB Amsterdam, The Netherlands

5 Tilburg University, Postbus 90153, 5000 LE Tilburg, The Netherlands

6 Foundation for Benchmarking Mental Health Care, Rembrandtlaan 46, 3723 BK Bilthoven, The Netherlands

7 University of Leiden, Postbus 9500, 2300 RA Leiden, The Netherlands

8 GGZ inGeest, A.J., Ernststraat 1187, 1081 HL Amsterdam, The Netherlands

from both perspectives and where patients are supported to participate actively in decision-making about treatment [6]. To participate in clinical decision-making, patients need access to information. Clinical feedback from outcome measurements in everyday practice, so-called Routine Outcome Monitoring (ROM), could be helpful in decision-making about treatment [7–9]. ROM is a personalised source of information and implies regular measurements of clinical outcomes during treatment. It provides feedback on the patients' progress [10, 11] and has the potential to be a useful tool to involve patients in their treatment process [7, 8, 12, 13]. Research has shown that ROM can enhance the communication between patients and clinicians, and, when patients and clinicians both are provided with this feedback information, ROM can have beneficial effects on mental health status, especially for patients who are not responding to treatment favourably [11, 13–16].

Hence, integrating systematic clinical feedback derived from ROM within a SDM framework could be considered a promising approach to improve the communication between patients and clinicians, encourage the empowerment of patients, and reduce decisional conflict (DC) [8]. DC refers to the degree to which patients were engaged in and felt comfortable about important clinical decisions [17–19]. In general, patients with less DC experience a better decision-making process. This leads to higher patient satisfaction with the clinician and can have positive effects on treatment engagement, well-being, and treatment outcome, whereas high DC has negative consequences in terms of decisional delay, regret, treatment dropout, worse clinical outcomes, and a higher intention to complain about treatment [17, 19–22]. Therefore, in the National Quality Improvement Collaborative for the clinical use of ROM, we invested in the implementation and evaluation of SDM using ROM (SDMR) in Dutch specialist mental health care. Although scientific knowledge about SDM [3, 5] and ROM [11, 13–16] as separate initiatives is expanding, no previous randomised-controlled trial (RCT) has investigated a combined SDMR approach in mental health care. As far as we know, only one pilot study [7] on the combination of SDM and ROM was conducted. This study showed the feasibility and utility of ROM as a personalised source of information while making shared decisions about treatment. To test the effects of this combined approach, an RCT was recommended. Here, we present the results of a multi-centre two-arm cluster RCT aimed to investigate the effects of this approach on DC, the working alliance between patients and clinicians, and treatment outcomes. We compared SDMR with Decision-Making As Usual (DMAU) and hypothesized that SDMR: (1) diminishes patients' perception of decisional conflict, (2) improves the working alliance between patients and clinicians, and (3) leads to better treatment outcomes.

Methods

Study design and participants

This study was a multi-centre two-arm match-paired cluster randomised-controlled trial (RCT) conducted in Dutch specialist mental healthcare in which the application of Shared Decision-Making using Routine Outcome Monitoring (SDMR) was compared with decision-making as usual (DMAU). 14 teams (7 intervention, 7 control) of 4 specialist mental health care organisations were recruited to participate in this trial. The design of this study has been described in more detail elsewhere [8].

SDMR was intended to be suitable for a broad population of patients treated in mental health care. Therefore, the trial involved teams treating patients with various diagnoses. The teams came from several catchment areas across the country (urban, semi-urban, and rural). To keep contamination to a minimum, we used a cluster design at the team level with pairs of two similar teams from the same mental health organisation being randomised.

Two complementary methods for recruiting patients were employed. First, clinicians of the intervention and control group consecutively invited each new patient to participate in the study. Second, in long-term treatment teams, patients were asked to participate at their annual treatment evaluation date. Patients were excluded from participating if they did not speak or read the Dutch language or were incapable of completing questionnaires because of cognitive functioning or an ongoing crisis. Patients enrolled after receiving face to face and written information about the study and after giving written informed consent.

Results are presented in accordance with the CONSORT statement for cluster randomised-controlled trials [23]. The trial was registered with The Netherlands Trial Registry (NTR5262).

Randomisation and blinding

After enrolment, random allocation of pairs of teams to either the experimental or control condition took place at cluster level by an independent statistician generating a random number (0 = DMAU, 1 = SDMR) by a SPSS syntax. The pairs of teams of the same mental health organisation treated a similar population of patients in the same geographical catchment area. Seven teams were allocated to the intervention group and seven teams to the control group. Due to the cluster randomisation and the nature of the intervention, blinding of the clinicians and patients was not possible. To reduce the risk of bias, data collection

was carried out by research assistants, independent of the research team and participating teams, with separate instruments, which were not part of the intervention.

Intervention

The RCT was part of a National Quality Improvement Collaborative (QIC), aimed to implement Routine Outcome Monitoring (ROM) in every day clinical practice [8, 9]. The intervention teams, which participated for a full year in the QIC program, implemented a model with five steps to apply SDMR as a personalised source of information [8]. An essential part of this intervention was the implementation of ROM, tailored to the patient group, in routine clinical practice [9]. In addition, prior to the study, the clinicians of the intervention teams underwent a 1-day training in applying SDMR in clinical practice. Over the course of this study, the clinicians of all the intervention teams received one central booster session, and additionally intervention teams organised their own local, regular supervision sessions. The researcher attended the supervision sessions of each team twice, aiming to monitor intervention integrity. A comprehensive description of the intervention can be found elsewhere [8, 9].

Data collection

Patients completed self-report questionnaires just after the start of SDMR (T0) and at 6-month follow-up (T1). Clinicians answered questions regarding their patients at 6 months (T1). Respondents received a link by email to complete the questionnaires and if necessary received one or more reminders by email, mail, or phone. If patients did not use Internet, they received paper questionnaires by mail.

Outcomes

Primary outcome

Decisional conflict was measured in patients using the revised, validated decisional conflict scale (DCS) [17], which was translated into Dutch [8, 18]. Each of the 16 items is scored from 0 (strongly agree) to 4 (strongly disagree). Besides a total score, the DCS includes five dimensions (information, support, clarification of values, certainty, and decision quality). Higher scores indicate more decisional conflict, which means that patients experience less information, less support, less clarification, less certainty, and poorer decision quality about decision-making. To calculate the total scale and scores of the five dimensions, the item scores were summed, divided by the number of items, and multiplied by 25. The scores thus range from 0 to 100 [17].

In general, the internal consistency of the total score and five dimensions of the Dutch version of the DCS calculated in this study population is good [18].

To compare patients' and clinicians' views on Decisional Conflict, additionally, DC was measured on a Visual Analogue (VAS) 10-point scale (item: 'to what extent do you agree with the decision taken?'), filled out at T1 by both patients and clinicians (regarding their patients).

Secondary outcomes

We used the following validated self-report questionnaires to measure the secondary outcomes: Working Alliance Inventory Short Form (WAI-S) [24], Outcome Questionnaire (OQ-45) [25], and Manchester Short Quality of Life Measurement (MANSA-VN-16) [26]. At T1, the WAI-S was filled out by both patients and clinicians. Measurement of treatment outcomes was tailored to the patient group. Therefore, the OQ-45, which measures symptom severity and functioning, was completed by patients in short-term treatment, and the MANSA-VN-16, which gives insight in the patients' perspective on the quality of life, by those in long-term treatment. Cronbach's alphas of these questionnaires calculated in this study population were: WAI-S $\alpha = 0.93$; MANSA-VN-16 $\alpha = 0.89$; OQ-45 $\alpha = 0.93$.

Process outcomes

The process outcomes gave information about the intervention integrity, which tests whether SDMR was properly applied in clinical practice. We tested this both in the experimental and in the control teams, because (elements of) SDMR have been adopted by many clinicians already. The SDM process was measured with the validated SDM-Q-9 [27], and the application of ROM with four questions about the clinical use of ROM feedback [8]. At T1, both questionnaires were completed by patients and clinicians. The calculated Cronbach's alphas in this study population were: SDM-Q-9 $\alpha \geq 0.91$ and ROM $\alpha \geq 0.77$.

Patients' characteristics

At baseline, socio-demographics (gender, age, and educational level) and clinical characteristics (setting, primary diagnoses, and locus of control) [28] were collected.

Statistical analysis

Analyses were based on the intention-to-treat principle. Socio-demographical and clinical characteristics were calculated using descriptive statistics. We also tested for possible differences in these characteristics between the two arms at T0 with independent samples t tests for the continuous and

Chi-square tests for the categorical variables [8]. To test for selective dropout before first or follow-up measurement, differences in baseline characteristics were assessed by logistic regression analyses. If participants had answered at least 80% of the items on a questionnaire, missing items were imputed with the mean value of the completed items. To take into account cluster effects and to handle missing data [29], the analyses of the primary and secondary outcomes and process parameters were performed using linear multilevel regression analysis (MLA) with three levels (teams, clinicians and patients). The analyses of the outcomes were based on the initial hypotheses performed with a two-tailed 0.05 significance level. We were also interested in the effect sizes (Cohen's *d*) of the differences between the two arms, which were calculated by dividing the between-group difference by the pooled SD. The thresholds for interpreting the effect sizes were: small 0.00–0.32, medium 0.33–0.55, and large ≥ 0.56 [30]. Furthermore, the differences in patients' and clinicians' views were tested. Finally, two post hoc analyses were performed for the primary outcome DC. First, we assessed potential effect modification by primary diagnoses and the degree of DC at T0. Interactions were tested for significance at a *p* value of 0.1. Second, we performed, in line with the initial hypotheses, per-protocol analyses by investigating the influence of intervention integrity on DC and the association between DC and treatment outcomes. MLA was performed in MLwiN 2.21 software. All other analyses were performed in SPSS 22.0.

Sample size

This study was powered to detect a medium-effect size of $d = 0.5$ on the primary outcome between intervention and control group. A sample size of 65 patients per arm was needed to obtain $\beta = 0.80$ power with alpha set at 0.05 (two-tailed). To correct for clustering, we used an expected intra-cluster correlation coefficient (ICC) of 0.05 and calculated [23, 31] that a sample size of 136 patients per arm was needed. Taking into account a dropout rate of 25%, we intended to include 182 patients per arm.

Results

Participant flow

Inclusion took place between October 2015 and July 2016. All measurements were completed in March 2017. 14 teams (7 intervention and 7 control) of 4 organisations for specialist mental health care and 55 clinicians participated in the trial. As shown in Fig. 1, 227 patients gave informed consent and 186 patients (94 intervention,

51%; 92 control, 49%) responded to the first measurement (T0). In a total 158 patients, 81 intervention (51%) and 77 control (49%) completed the follow-up measurement at 6 months (T1). The range of participating patients per team was 4 to 28 (mean 13.3), and per clinician, the range was 1 to 11 patients (mean 3.38). The number of patients was in balance between the pairs of teams in the two arms. Clinicians completed an assessment at 6 months for 173 patients: 90 (52%) from the intervention and 83 (48%) from the control group. The dropout rate of patients between inclusion and T0 was 18%, and between T0 and T1, it was 11%. Numbers of dropout were not significantly different between the intervention and control group. In addition, the dropout rate was not significantly associated with any socio-demographic and most clinical characteristics. Only for locus of control was a significant difference found. Participants who dropped out between first and follow-up measurement reported significantly higher locus of control ($p = 0.034$) compared to patients who completed the follow-up measurement. Reasons for dropout were: (1) withdrawal of informed consent (57%), (2) inability to complete questionnaires (23%), because of decreased cognitive capacity, current crisis, or death, and (3) patients not responding to reminders (20%).

Patients' characteristics

Table 1 shows a complete overview of the socio-demographics (gender, age, and educational level) and clinical characteristics (setting, primary diagnosis, and locus of control) of the study population, which were balanced between the two arms. Patients with the primary diagnoses depressive (26%) and personality disorders (20%) were relatively the largest groups, followed by, respectively, psychotic (17%), anxiety (15%), bipolar (13%), and other* (9%) disorders. Patients' characteristics showed the variety of participating patients receiving different kinds of long-term and short-term treatment. All patients were treated in outpatient settings, but followed different types of treatment (for example: different kinds of individual or group psychotherapy, medication treatment, or support and guidance in daily life given by nurses). We could not extract information from electronic records at patient level about who exactly received which kinds of treatment. Because of these varieties in patients and treatment characteristics, patients who participated in this trial also had to make decisions about various issues such as: diagnostic research, type of therapy, choices in medication treatment, personal treatment goals, or daytime activities.
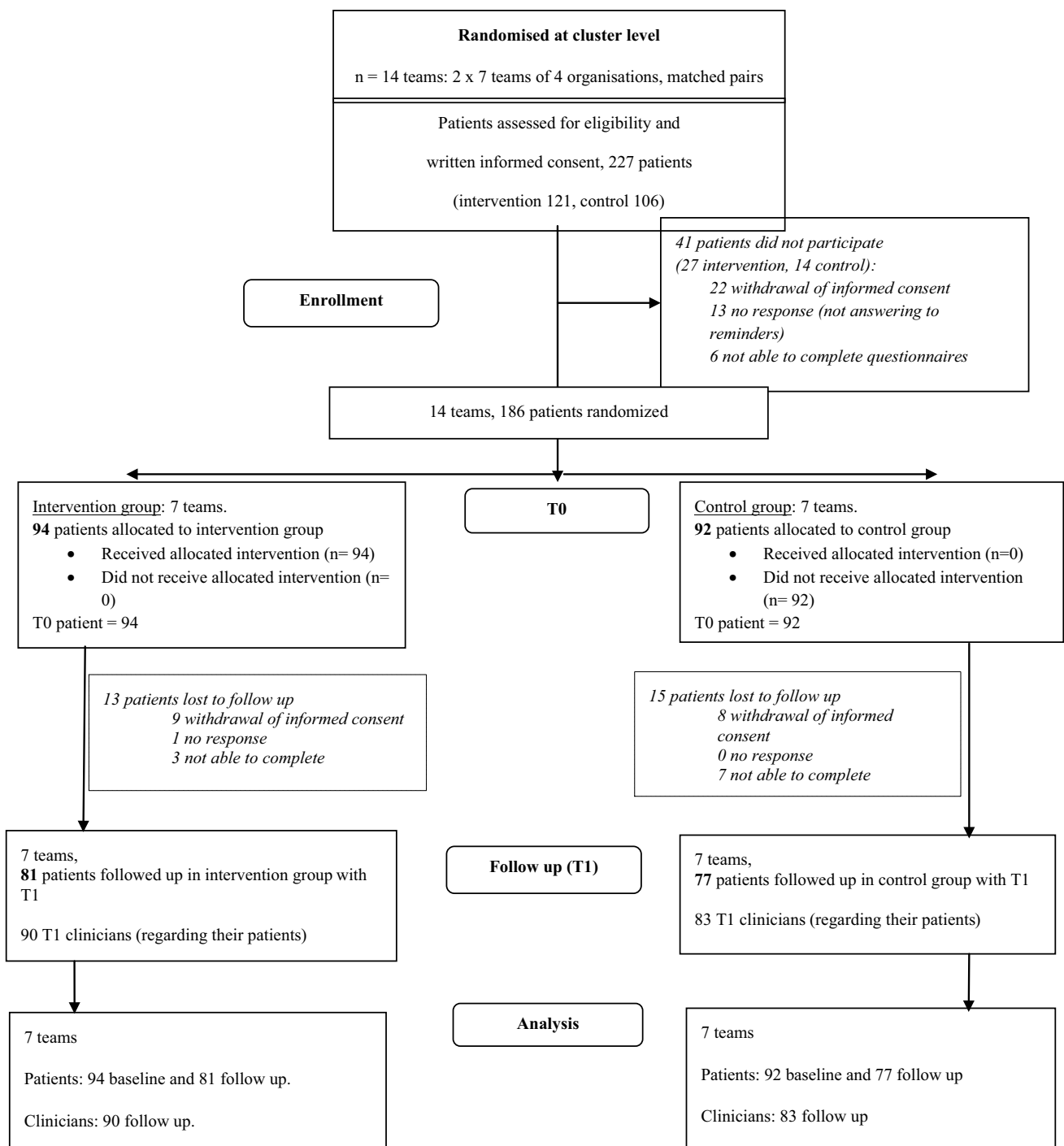
Fig. 1 Flow diagram RCT shared decision-making and ROM (Consort 2010)

Effectiveness of the intervention

In the intention-to-treat analyses, no significant differences between SDMR and DMAU on the primary outcome DC were found (Table 2), either for the total DC scale, or on the subdomains. We only found cluster effects within clinicians (ICC 0.2) and did not find evidence for clustering of the results at team level (ICC 0.0), which was also the unit of randomisation. Regarding the DC VAS scale, no significant differences were found for the patient-reported version. Only, for the clinician, completed version was a difference found. Compared to the control group, clinicians in the intervention group reported on this VAS scale significantly more 'agreement with the decision taken' ($\beta = 0.710$,

Table 1 Participants' characteristics

		Total study population ($n = 186$)	Intervention group ($n = 94$)	Control group ($n = 92$)
Socio-demographics				
Gender	Female	111 (59.7%)	59 (62.8%)	52 (56.5%)
	Male	75 (40.3%)	35 (37.2%)	40 (43.5%)
Age	Mean age (sd) (min, max)	47.2 (18.0) (18–83 years)	47.2 (18.0) (18–80 years)	47.2 (18.2) (18–83 years)
Educational level	Primary school or Lower secondary education (low)	68 (36.6%)	33 (35.1%)	35 (38.0%)
	Higher secondary or intermediate vocational education (middle)	87 (46.8%)	45 (47.9%)	42 (45.7%)
	Higher vocational education or university/Bachelor's or Master's degree (high)	31 (16.7%)	16 (17.0%)	15 (16.3%)
Clinical characteristics				
Setting	Cure (short term)	100 (53.8%)	50 (53.2%)	50 (54.3%)
	Severe mental illness (long term)	86 (46.2%)	44 (46.8%)	42 (45.7%)
Primary diagnosis	Depressive disorder	48 (25.8%)	26 (27.6%)	22 (23.9%)
	Personality disorder	38 (20.4%)	13 (13.8%)	25 (27.2%)
	Psychotic disorder	32 (17.2%)	20 (21.3%)	12 (13.0%)
	Anxiety disorder	27 (14.5%)	11 (11.7%)	16 (17.4%)
	Bipolar disorder	24 (12.9%)	15 (16.0%)	9 (9.8%)
	Other disorders*	17 (9.1%)	9 (9.6%)	8 (8.7%)
Locus of control	Q1	Q1 11.0	Q1 10.0	Q1 11.0
	Q2, median	Q2 13.0	Q2 13.5	Q2 13.0
	Q3	Q3 16.0	Q3 16.0	Q3 17.0
	Minimum	5	5	5
	Maximum	25	23	25

*Other disorders: dissociative disorder NAO, dementia NAO, mood disorder by alcohol, undifferentiated somatoform disorder, hypochondria, and childhood disorder

No significant differences were found between the two arms in socio-demographics and clinical characteristics' ($p < 0.05$)

$p = 0.018$, 95%CI: 0.122–1.298) with a medium-effect size ($d = 0.45$). In addition, we found no significant effects of SDMR on patients' reported secondary outcomes: working alliance, clinical outcome, and quality of life. According to clinicians, the working alliance was better in the intervention group as compared to the control group ($\beta = 0.254$, $p = 0.026$ 95%CI: 0.031–0,477; $d = 0.45$). Finally, although the application of SDM reported by patients and clinicians did not differ between conditions, both patients and clinicians experienced a significantly better usage of ROM in the intervention group compared to the control group (patients $\beta = 10.980$, $p = 0.013$, 95%CI: 2.289–19.671; clinicians: $\beta = 25.558$, $p = 0.000$, 95%CI: 17.930-33.186). The effect sizes were medium (patients: $d = 0.39$) to large (clinicians: $d = 1.31$).

When comparing patients' with clinicians' reported measurements, we found, irrespective of the condition, that clinicians, compared to patients, scored significantly higher (more positive) on the DC VAS scale (mean$_{\text{diff}}$. 1.55, sd 2.96, $p = 0.000$, 95%CI: 1.069–2.028), WAI-S (mean$_{\text{diff}}$. 0.39, sd 0.98, $p = 0.000$, 95%CI: 0.240–0.539), SDMQ9

(mean$_{\text{diff}}$. 7.79, sd 32.46, $p = 0.004$, 95%CI: 2.531–13.040), and ROM questions (mean$_{\text{diff}}$. 10.64, sd 35.19, $p = 0.000$, 95%CI: 4.941–16.33).

Subgroup analyses

Post hoc analyses for the primary outcome DC (DCS total scale $_{\text{change T1–T0}}$) showed significant interaction effects between primary diagnosis and trial condition, which means that the effect of SDMR on the primary outcome was significantly influenced by the primary diagnosis as effect modifier. Table 3 demonstrates the intervention effects for each group of primary diagnoses. Patients with the primary diagnoses' depression and anxiety demonstrated a reduced degree of DC in the intervention group at T1, whereas the other diagnosis groups showed an increased level of DC at follow-up. For depressed patients, SDMR had a large effect ($d = -0.69$) and was significantly more effective than DMAU ($\beta = -11.173$, $p = 0.047$, 95%CI: -22.176 to -0.170). Furthermore, the level of experienced DC at T0 was also an effect modifier for the DCS total change score (Table 3).

Table 2 Differences between intervention and control group on the primary and secondary outcomes and process parameters

Primary outcome decisional conflict

	T0	T1	β	95% CI	p value	Effect size
	Mean (sd) T0, T1		Change score T1−T0			
Total score	C 39.79 (15·37)	C 33.18 (16.52)	1.234	− 5.457 to 7.925	0.717	0.07
	I 37.78 (18·99)	I 33.03 (19.32)				
Information	C 44.51 (19·85)	C 35.61 (19.99)	1.353	− 5.875 to 8.581	0.713	0.06
	I 41.58 (22·56)	I 33.64 (22.22)				
Support	C 36.36 (19·62)	C 31.93 (17.70)	2.942	− 3.181 to 9.065	0.346	0.15
	I 36.70 (20·68)	I 34.05 (22.29)				
Clarification	C 39.26 (19.12)	C 32.90 (19.49)	− 0.788	− 9.771 to 8.195	0.862	− 0.03
	I 38.12 (21.57)	I 32.10 (21.61)				
Certainty	C 46.15 (20.84)	C 37.88 (22.48)	3.962	− 3.533 to 11.457	0.300	0.16
	I 41.58 (24.13)	I 36.52 (24.24)				
Quality	C 35.00 (17.01)	C 28.98 (18.08)	4.006	− 2.268 to 10.280	0.211	0.20
	I 32.91 (19.98)	I 29.92 (20.40)				
	Mean (sd) T1		T1			
VAS DC patient	C 7.40 (2.47)		− 0.031	− 0.819 to 0.757	0.938	− 0.012
	I 7.37 (2.60)					
VAS DC clinician	C 8.45 (1.79)		0.710	0.122 to 1.298	0.018	0.45
	I 9.17 (1.32)					

Secondary outcomes
Working alliance (WAV), clinical outcome (OQ-45), quality of life (MANSA-VN-16)

	T0	T1	β	95% CI	p value	effect size
	Mean (sd) T0, T1		Change score T1–T0			
WAV patient	C 3.53 (0.81)	C 3.57 (0.89)	− 0.020	− 0.257 to 0.217	0.869	− 0.03
Total score	I 3.48 (0.88)	I 3.54 (0.90)				
	Mean $_{T1}$		T1			
WAV clinician	C 3.75 (0.53)		0.254	0.031 to 0.477	0.026	0.45
Total score	I 4.00 (0.60)					
OQ-45	C 78.06 (22.52)	C 71.69 (24.73)	− 0.964	− 9.702 to 7.774	0.828	− 0.05
Total score	I 83.49 (27.52)	I 78.07 (31.08)				
MANSA-VN-16	C 4.25 (1.06)	C 4.35 (0.94)	− 0.234	− 0.491 to 0.022	0.074	− 0.42
Total score	I 4.43 (0.99)	I 4.30 (0.94)				

Process parameters
SDM (SDMQ9) and ROM process

	Mean $_{T1}$	T1				
SDMQ9 patient	C 65.22 (23.94)	− 0.834	− 8.405 to 6.737	0.828	− 0.03
	I 64.39 (24.88)				
SDMQ9 clinician	C 69.61 (18.92)	5,544	− 1.349 to 12.437	0.115	0.31
	I 75.78 (16.22)				
ROM patient	C 48.90 (28.30)	10.980	2.289 to 19.671	0.013	0.39
	I 59.88 (27.78)				
ROM clinician	C 49.70 (22.56)	25.558	17.930 to 33.186	0.000	1.31
	I 76.44 (16.01)				

C control group, *I* intervention Group, *CI* confidence interval

Table 3 Results of subgroup analyses: differences in change scores (T1–T0) on the primary outcome DC in the intervention group for diagnosis groups and level of DC at T0

Primary outcome decisional conflict (total score)

	T0	T1	β	95% CI	p value	Effect size
	Mean (sd) $_{T0, T1}$		Change score $_{T1 - T0}$			
Depression	C 36.01 (12.53)	C 39.58 (16.10)	− 11.173	− 22.176 to − 0.170	0.047	− 0.69
	I 38.47 (16.55)	I 31.96 (22.00)				
Anxiety disorder	C 38.57 (11.51)	C 24.86 (11.28)	− 4.003	− 18.297 to 10.291	0.583	− 0.30
	I 46.31 (22.62)	I 33.52 (15.83)				
Psychotic disorder	C 38.80 (15.13)	C 28.75 (9.29)	13.206	− 1.545 to 27.957	0.079	0.97
	I 29.12 (16.87)	I 27.54 (16.08)				
Personality disorder	C 44.06 (15.84)	C 35.81 (20.23)	3.478	− 8.750 to 15.706	0.577	0.14
	I 47.60 (18.17)	I 44.89 (22.73)				
Bipolar disorder	C 33.98 (16.15)	C 23.18 (15.16)	11.055	− 5.433 to 27.543	0.189	0.65
	I 39.58 (19.25)	I 35.55 (14.90)				
Rest category	C 46.09 (24.32)	C 35.35 (13.15)	6.575	− 10.136 to 23.286	0.440	0.59
	I 27.43 (17.44)	I 26.37 (19.77)				
Low level of DC at T0	C 26.83 (6.27)	C 27.71 (15.13)	− 1.777	− 9.807 to 6.253	0.665	− 0.12
	I 24.43 (9.83)	I 24.20 (13.99)				
High level of DC at T0	C 50.64 (11.89)	C 37.50 (16.45)	1.203	− 6.813 to 9.219	0.768	0.06
	I 55.03 (13.10)	I 44.64 (19.36)				

Compared to the control group, patients in the intervention group with a lower level of DC at T0 reported more reduction in DC, while patients who experienced a high level of DC at T0 showed less reduction in DC at T1.

Associations between process and outcome parameters

As part of the post hoc analyses, we explored the association between the process parameters SDMQ9 and ROM with the primary outcome DC (Table 4). We found that a better application of SDM and ROM, according to patients, led to significantly less DC (SDM $\beta = − 0.217$, $p = 0.000$, 95%CI: − 0.321 to − 0.113; ROM $\beta= − 0.111$, $p = 0.021$,

95%CI: -0.205 to − 0.017). The effect sizes were small (ROM $d = − 0.32$)-to-medium (SDM $d = − 0.45$). When clinicians reported a better SDM and ROM application, we did not find significantly different DC scores in patients. Thus, in contrast to the clinicians' reported level of SDM and ROM, when patients experienced a better application of SDM and ROM, this was associated with significantly decreased DC scores. Finally, we explored the influence of DC on treatment outcomes (Table 4). We found significant associations (medium-effect sizes) between reduced DC and better clinical outcomes ($\beta = 0.344$, $p = 0.001$, 95% CI 0.134 to 0.554, $d = 0.40$,) and a higher quality of life ($\beta = − 0.011$, $p = 0.006$, 95% CI − 0.019 to − 0.0032, $d = − 0.54$).

Table 4 Associations of process parameters with the primary outcome DC and the influence of DC change score on treatment outcomes

		β	95% CI	p value	Effect size
Associations between process parameters and primary outcome					
Decisional conflict	SDMQ9 patient $_{T1}$	− 0.217	− 0.321 to − 0.113	0.000	− 0.45
Total score $_{change\ T1–T0}$	SDMQ9 clinician $_{T1}$	0.014	− 0.164 to 0.192	0.878	0.23
	ROM patient $_{T1}$	− 0.111	− 0.205 to − 0.017	0.021	− 0.32
	ROM clinician $_{T1}$	− 0.044	− 0.179 to 0.091	0.523	0.00
Associations between primary outcome and treatment outcomes					
OQ-45	Decisional conflict	0.344	0.134 to 0.554	0.001	0.40
Total score $_{change\ T1–T0}$	Total score $_{change\ T1 – T0}$				
MANSA-VN-16	Decisional conflict	− 0.011	− 0.019 to − 0.0032	0.006	− 0.54
Total score $_{change\ T1–T0}$	Total score $_{change\ T1–T0}$				

Discussion

This study presents the results of a cluster randomised-controlled trial, testing the application of Shared Decision-Making using Routine Outcome Monitoring (SDMR) in specialist mental health care. Compared to decision-making as usual (DMAU), we found no significant influence of SDMR on the primary outcome reduction of Decisional Conflict (DC) reported by patients. Furthermore, there were no significant differences in the patients' rated secondary outcomes (working alliance, clinical outcome, and quality of life) between the two arms. In contrast to the patients' rated outcome parameters, the clinicians in the intervention group reported a significantly better working alliance and experienced more agreement with the clinical decisions that were taken. Subgroup analyses showed that that effect of the intervention was significantly modified by diagnosis. Depressed patients reported a significant and large effect of SDMR on reduction of DC. Looking at important process parameters, both patients and clinicians in the intervention group reported a significantly better application of ROM. Patients reported medium effects, while clinicians reported very large effects. The SDM process rated by clinicians and patients did not differ between the groups. In general, irrespective of the condition, clinicians' views about the outcome and process parameters were significantly more positive than patients' rated parameters. Exploring the influence of the process parameters on DC demonstrated that a better implementation of SDMR was associated with significantly less DC experienced by patients. Thus, when, according to patients, SDMR was applied well, patients reported a reduction of DC, which had a positive influence on clinical outcomes and quality of life.

Interpretation and clinical implications

To our knowledge, this is the first RCT to rigorously test the effectiveness of a combined SDM with ROM initiative in mental health care. In contrast with our hypotheses and previous research [1–5, 11, 13–16], we found no significant differences between SDMR and DMAU regarding the primary and secondary outcomes reported by patients. Interestingly, the results of the clinicians on DC and working alliance in the intervention group were significantly more positive compared to the control group. A possible explanation for the lack of differences between the conditions on the patients' rated outcome parameters is that SDMR was not optimally implemented. First, this could be evidenced by the fact that, although clinicians in the intervention group reported change, patients did not notice

much difference. A second argument is that, despite the investment in training, booster, and supervision sessions, the application of SDM, reported by patients and clinicians, did not differ between the two arms. Only the usage of ROM was significantly better in the intervention group according to both patients and clinicians. When looking at the association between the process parameters and the primary outcome, we found, in line with our hypotheses and previous research [1–5, 11, 13–16], that a better application of SDMR led to less DC. Thus, if clinicians were able to adequately apply SDMR during treatment consultations with their patients, this approach resulted in positive effects for patients. Less DC was, in turn, associated with better clinical outcomes and better quality of life.

Moreover, the subgroup analyses indicated that the effect of the intervention differed between subgroups of patients, with a significantly positive effect for depressed patients. This is in line with previous research which revealed that shared decision-making had a positive impact on depressed patients in terms of satisfaction and treatment adherence [3]. Previous work also demonstrated that feedback during treatment led to better outcomes in patients with mood disorders [16, 32]. This may imply that the intervention was too generic for the variety of patients treated in specialist mental health care. In designing the study, we hoped to be able to arrive at a generic method to improve SDM across mental health, and we approached teams treating very different types of patients. However, tailoring the intervention to the patient group and its decision topics may be necessary to improve results. Previous work has shown that the success of SDM depends on the patients' ability to participate actively [33, 34]. A limitation of our intervention is that we focused our training on clinicians. Future interventions should targeting both patients and clinicians [33, 34].

Strengths and limitations

Our study had various strengths. First, due to the matched pair randomisation between similar teams, the risk of confounding was reduced. This was evidenced by the fact that no differences were found in patients' characteristics and dropout rates between the two arms. Second, because this study was conducted in real-world clinical practice with a heterogeneous patient population, including various ages, diagnostic groups, and treatment settings, the study results were generalizable to a large group of mental health care teams, and thus had good external validity. Because of this heterogeneity, we were also able to investigate differences in effects for subgroups of patients. Third, the independent data collection with separate research instruments apart from the intervention prevented socially desirable scoring of patients and undesired influence of the research team or clinicians on

the results. In addition, it enhanced the uniformity and quality of the collected data. Furthermore, at T1, we had insight into both patients' and clinicians' views on the application of SDMR. Finally, we used multilevel analyses to correct for the variability of the results at clinician level.

This study also had a number of limitations. Despite the extra efforts (recruiting an additional pair of two teams, checking if all eligible patients were asked to participate in the study and extending the inclusion period by a maximum of 5 months), the study population of 186 patients with a first measurement was smaller than intended. Reasons for this smaller population were difficulties in recruiting patients due to the lower inflow of patients receiving treatment from the participating teams and fewer patients who were willing to participate in the study. Because of the end date of the national collaborative and its funding, we could not extend the inclusion period further. Although the estimated sample size was not reached, the study population of 186 patients, with more than 90 patients per arm and the balance between arms in patients' characteristics and dropout rates, proved sufficient to detect medium-effect sizes. Looking at the effect sizes of the primary analyses, these were extremely small, suggesting that power was not a deciding factor in our negative results of the primary outcome. Furthermore, although we adopted a cluster randomised-controlled design to reduce the risk of contaminating the two conditions, contamination cannot be ruled out altogether and may have weakened the effects found. Finally, treatment outcomes were evaluated, tailored to the patient groups, with two different measurement instruments. Patients who received short-term treatment filled out the OQ-45, which measures symptom severity and functioning. Patients who followed long-term treatment completed the MANSA-VN-16 which measured quality of life. Therefore, we have partially insight into the effects of SDMR on symptom severity, merely for patients who followed short-term treatment.

Conclusions

To our knowledge, this is the first randomised-controlled trial (RCT) which investigated the combination of Shared Decision-Making using Routine Outcome Monitoring (SDMR) in specialist mental health care.

In general, this study showed no effects of SDMR on primary and secondary outcomes rated by patients. However, subgroup analyses revealed a positive effect of SDMR on DC for depressed patients. According to clinicians, SDMR has a positive impact on DC and the working alliance. Furthermore, in the intervention group, both patients and clinicians reported a better ROM application. Finally, taking participants of both arms into account, a higher level of SDMR led to less DC, which, in turn, had a positive influence on treatment outcomes. Although this is a negative trial, the results are encouraging for further investments in the development, implementation, and investigation of this approach in clinical practice.

Compliance with ethical standards

References

1. Malm U, Ivarsson B, Allebeck P, Falloon IRH (2003) Integrated care in schizophrenia: a 2-year randomised controlled study of two community-based treatment programs. Acta Psychiatr Scand 107:415–423
2. Clever SL, Ford DE, Rubenstein LV, Rost KM, Meredith LS, Sherbourne CD, Wang NY, Arbelaez JJ, Cooper LA (2006) Primary care patients' involvement in decision-making is associated with improvement in depression. Med Care 44:398–405
3. Patel SR (2008) Recent advances in Shared Decision Making for Mental Health. Curr Opin Psychiatry 21:606–612
4. Westermann GMA, Verheij F, Winkens B, Verhulst FC, Van Oort FVA (2013) Structured shared decision-making using dialogue and visualization: a randomised controlled trial. Patient Educ Couns 90:74–81
5. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt A, Thomson R, Trevena L (2017) Decision aids for people facing health treatment or screening decisions (review). Cochrane Database Syst Rev. https://doi.org/10.1002/14651858.CD001431.pub5

6. Elwyn G, Forsch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P (2012) Shared Decision Making: a model for clinical practice. J Gen Intern Med 27:1361–1367

7. Van der Feltz-Cornelis C, Andrea H, Kessels E, Duivenvoorden H, Biemans H, Metz M (2014) Does routine outcome monitoring have a promising future? An investigation into the use of shared decision making combined with ROM for patients with a combination of physicial and psychiatric symptoms. Tijdschr Psychiatr 56:375–384

8. Metz MJ, Franx GC, Veerbeek MA, de Beurs de E, van, der Feltz van der-Cornelis CM, Beekman ATF (2015) Shared decision making in mental health care using Routine Outcome Monitoring as a source of information: a cluster randomised controlled trial. BMC Psychiatry 15:313–312

9. Metz MJ, Veerbeek MA, Franx GC, Van der Feltz-Cornelis CM, De Beurs E, Beekman ATF (2017) A National Quality Improvement Collaborative for the clinical use of outcome measurement in specialised mental health care: results from a parallel group design and a nested cluster randomised controlled trial. BJPsych Open 3:106–112

10. De Beurs E, Den Hollander-Gijsman ME, Van Rood YR, Van der Wee NJ, Giltay EJ, Van Noorden MS, Van der Lem R, Van Fenema E, Zitman FG (2011) Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. Clin Psychol Psychother 18:1–12

11. Carlier IVE, Meuldijk D, Van Vliet IM, Van Fenema EM, Van der Wee NJA, Zitman FG (2012) Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. J of Eval Clin Pract 18:104–110

12. Eisen SV, Dickey B, Sederer LI (2000) A self-report symptom and problem rating scale to increase inpatients' involvement in treatment. Psychiatr Serv 51:349–353

13. Gondek D, Edbrooke-Childs J, Fink E, Deighton J, Wolpert M (2016) Feedback from outcome measures and treatment effectiveness, treatment efficiency, and collaborative practice: a systematic review. Adm Policy Ment Health 43:325–343

14. Knaup C, Koesters M, Schoefer D, Becker T, Puschner B (2009) Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. Br J Psychiatry 195:15–22

15. Shimokawa K, Lambert MJ, Smart DW (2010) Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. J Consult Clin Psychol 78:298–311

16. Guo T, Xiang YT, Xiao L, Hu CQ, Chiu HF, Ungvari GS, Correll CU, Lai KY, Feng L, Geng Y, Feng Y, Wang G (2015) Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. Am J Psychiatry 172:1004–1013

17. O'Connor AM (2010) User-Manual-Decisional Conflict Scale (16 item statement format). Hospital Research Institute, Ottawa. https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_decisional_conflict.pdf. Accessed 10 February 2015

18. Metz MJ, Veerbeek MA, Van der Feltz-Cornelis CM, De Beurs E, Beekman ATF (2018) Decisional conflict in mental health care: a cross-sectional study. Soc Psychiatry Psychiatr Epidemiol 53:161–169

19. Hölzel P, Kriston L, Härter M (2013) Patient preference for involvement, experienced involvement, decisional conflict, and satisfaction with physician: a structural equation model test. BMC Health Serv Res 13:231–210

20. Katapodi MC, Munro ML, Pierce PF, Williams RA (2011) Testing of the decisional conflict scale: genetic testing hereditary breast, ovarian cancer. Nurs Res 60:368–310

21. 21 Des Cormiers A, Légaré F, Simard S, Boulet LP (2015) Decisional conflict in asthma patients: a cross sectional study. J Asthma 52:1084–1088

22. 22 Thompson-Leduc P, Turcotte S, Labrecque M, Légaré F (2016) Prevalence of clinically significant decisional conflict: an analysis of five studies on decision-making in primary care. BMJ. https://doi.org/10.1136/bmjopen-2016-011490

23. Campbell MK, Piaggio G, Elbourne DR, Altmann DG (2012) Consort 2010 statement: extension to cluster randomised trails. BMJ 345:1–21

24. Stinckens N, Ulburghs A, Claes L (2009) De werkalliantie als sleutelelement in het therapiegebeuren. Meting met behulp van de WAV-12, de Nederlandstalige verkorte versie van de Working Alliance Inventory. TKP 1:44–60

25. De Jong K, Nugter MA, Lambert MJ, Burlingame GM (2008) Handleiding voor afname en scoring van de Outcome Questionnaire OQ-45.2. OQ Measures, Salt Lake City

26. Van Nieuwenhuizen Ch, Janssen-de Ruijter EAW, Nugter MA (2017) MANSA: Manchester Short Quality of Life Handleiding

27. Rodenburg-Vandenbussche S, Pieterse AH, Kroonenberg PM, Scholl I, Van der Weijden T, Luyten GP, Kruitwagen RF, Den Ouden H, Carlier IV, Van Vliet IM, Zitman FG, Stiggelbout AM (2015) Dutch translation and psychometric testing of the 9-item Shared Decision Making Questionnaire (SDM-Q-9) and Shared Decision Making Questionnaire-Physician version (SDM-Q-Doc) in primary and secondary care. PLoS One. https://doi.org/10.1371/journal.pone.0132158

28. Pearlin LI, Lieberman MA (1981) The stress process. J Health Soc Behav 22:337–356

29. Twisk JWR (2006) Chap. 6 Multilevel analysis in longitudinal studies. In: Twisk JWR Applied multilevel analysis, 6th printing. Cambridge university press, Cambridge, pp 86–107

30. Lipsey MW, Wilson DB (1993) The efficacy of psychological, educational and behavioural treatment: confirmation from meta-analysis. Am Psychol 48:1181–1209

31. Killip S, Mahfoud Z, Pearce K (2004) What is an Intracluster correlation coefficient? Crucial concepts for primary care researchers. Ann Fam Med 2:204–208

32. Janse PD, De Jong K, Van Dijk MK, Hutschemaekers GJM, Verbraak MJPM (2016) Improving the efficiency of cognitive-behavioural therapy by using formal client feedback. Psychother Res 27:525–538

33. Joseph-Williams N, Elwyn G, Edwards A (2014) Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns 94:291–309

34. Légaré F, Turcotte S, Stacey D, Ratté S, Kryworuchko J, Graham ID (2012) Patients' perceptions of sharing in decisions: a systematic review of interventions to enhance shared decision making in routine clinical practice. Patient 5:1–19