Is fidelity to motivational interviewing associated with alcohol outcomes in treatment-seeking 60+ year-old citizens?

Kramer Schmidt, Lotte; Moyers, Theresa B.; Nielsen, Anette Søgaard; Andersen, Kjeld

Published in:
Journal of Substance Abuse Treatment

DOI:
10.1016/j.jsat.2019.03.004

Publication date:
2019

Document version
Final published version

Document license
CC BY-NC-ND

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
Is fidelity to motivational interviewing associated with alcohol outcomes in treatment-seeking 60+ year-old citizens?

Lotte Kramer Schmidt, Theresa B. Moyers, Anette Seggaard Nielsen, Kjeld Andersen

ABSTRACT

Background: Part of the variability in treatment outcomes for Motivational Interviewing (MI) may be explained by differences in the fidelity to MI. The Motivational Interviewing Treatment Integrity manual version 4 (MITI 4) is an improved measure of fidelity to elements of MI. It is not known whether the fidelity to MI, as measured by the MITI 4, is related to treatment outcome.

Objectives: To examine whether fidelity to MI is associated with alcohol use outcomes – predictive validity of the MITI 4.

Method: Twenty percent of the recorded sessions at the Danish sites of the Elderly Study were randomly drawn and coded for fidelity to MI with the MITI 4. The Elderly Study was an international, randomized controlled trial, in which people 60 years or older with Alcohol Use Disorders received either four weeks of Motivational Enhancement Therapy (MET) or four weeks of MET combined with up to eight additional sessions of the Community Reinforcement Approach-Senior (MET+CRA-S). Elements of MI and summary scores of the MITI 4 were used as predictors in a mixed effects regression analysis. Treatment outcomes were use of alcohol and consequences of drinking at 26-weeks follow-up.

Results: In total, 423 sessions representing 238 participants were randomly drawn and coded for fidelity to MI. Mean values of the treatment elements indicated high fidelity to MI, with higher fidelity to MI in the MET sessions, as compared to CRA-S sessions. None of the predictors in the multilevel model analyses were associated with outcome at follow-up. Exploratory analysis indicated reverse associations between one measure of MI-fidelity and drinking outcomes in the combined treatment (CRA-S).

Conclusion: The fidelity of the MI intervention, received by participants in this study, did not predict better treatment outcomes. MI may be less effective in populations which are already committed to change behavior. As expected and validating for the MITI 4, fidelity to MI-elements was lower in the combination of MI with other treatment approaches. Additionally, the timing of MI in these combined settings might be important for effectiveness.

1. Introduction

1.1. Background

Treatment fidelity refers to how well a treatment is implemented (Perepletchikova, 2011). Although Motivational Interviewing (MI) has demonstrated efficacy in the treatment of Alcohol Use Disorders (AUD) (Smedslund et al., 2011), there is considerable variation in effect sizes, as well as negative trials. This has raised the concern that the fidelity, e.g. implementation, of MI may be poor – that MI in theory and MI in clinical practice may be different things and that this in turn could be responsible for confusing results of MI clinical trials (Miller and Rollnick, 2014).

1.2. Fidelity and outcome

Does fidelity to motivational interviewing influence the outcomes of participants receiving MI? A definitive answer to this question would require a manipulation of the fidelity to MI as it is offered in a clinical trial – an unlikely and expensive proposition. Instead, we evaluate both the fidelity to MI and the outcomes of clients in clinical trials, to examine any relationship between them. If MI works as hypothesized, by conveying a relational and technical component (Miller and Rose, 2009), there should be clear associations between the fidelity of these elements and client outcomes. Although MI therapist behaviors are commonly linked to outcomes in studies of MI (Borsari et al., 2015; Gaume et al., 2014; Gaume, Heather, Tober, and McCambridge, 2018; Grodensky et al., 2017), meta-analyses of this hypothesized association have not detected direct associations between MI elements and outcome (Magill et al., 2018; Pace et al., 2017). Indications were found.
however, for **indirect** pathways of associations between MI-elements and outcome; that is, through measures of client change talk or client sustain talk leading to change outcomes. The ability to measure the therapist’s attention to the client’s language about change, therefore, is emerging as a critical element of MI competence and one which is difficult to measure with objective behavioral ratings systems. This is because strategic responding to change and sustain talk often requires a knowledge of the therapist’s intent (or strategy), which is difficult to capture in discrete behaviors.

1.3. The motivational interviewing treatment integrity code (MITI)

The Motivational Interviewing Treatment Integrity Code (MITI) (Moyers, Rowell, Manuel, Ernst, and Houck, 2016) is a frequently used tool to measure fidelity of MI. The scores of the MITI is a measure of fidelity to MI through measures of frequency and competence of specific elements of MI; moreover, it measures if the therapist is conveying elements that are not recommended when performing MI. The MITI measures fidelity in two dimensions: 1) a quantitative, logical counting of therapist behaviors explicitly described with decisional rules, and 2) a global assessment of competence with anchors on a five- point Likert scale assessing “the gestalt” of the therapist (Moyers, Manuel, and Ernst, 2014). Please find links to The Motivational Interviewing Treatment Integrity Coding Manual 4.2.1 (MITI 4) at the end of the reference list. Previous versions of the MITI have demonstrated acceptable validity and reliability (Forsberg, Berman, Kalilmen, Hermanssion, and Helgason, 2008; Forsberg, Kalilmen, Hermanssion, Berman, and Helgason, 2007; Moyers, Martin, Manuel, Hendrickson, and Miller, 2005; Pierson et al., 2007) in assessing the fidelity of the therapist’s use of MI. The fourth revision of the MITI incorporated the empirical literature, indicating the importance of the therapist’s role in both encouraging change, favoring language from the client (Change Talk) and decreasing language opposing change (Sustain Talk) in MI treatment sessions. The new measures of the MITI 4 has acceptable reliability (Moyers et al., 2016; Copeland et al., 2018; Kramer Schmidt, Andersen, Nielsen, and Moyers, 2019; Owens, Rowell, and Moyers, 2017; Serrano et al., 2017) and validity when comparing them to the golden standard, the MISC (Motivational Interviewing Skills Code) (Moyers et al., 2016).

Early evidence indicates that the MITI 4 measures fidelity to elements of the MI process: Owens et al. (2017), assessed associations between MITI 4 global ratings and motivation and confidence to decrease drug use by the end of treatment among incarcerated individuals. In particular, significant associations were found between the technical score of the MITI 4 and motivation to decrease drug use by the end of treatment (Owens et al., 2017). The technical score is based on how well the therapist is cultivating and evoking client change talk, and simultaneously softening client sustain talk (Moyers et al., 2014). Although this link between MITI 4 scores and client motivation is promising, predictive validity for alcohol use outcomes has not yet been explored.

If accurate measures of elements of MI are predictive of client outcome, it will allow for powerful explorations of how MI affects change of alcohol use. Furthermore, insight in the relation between elements of MI and patient outcome may lead to more efficient training and treatment delivery. Equally important, MI itself may be improved if careful analysis can lead to elimination of superstitious elements that do not lead to better outcome, but only imply costs for the therapist and client.

1.4. MI combined with other therapies

The effects of MI could be even better in combination with other treatment approaches (Hettema, Steele, and Miller, 2005; Spoelstra, Schueler, Hilton, and Ridouen, 2015; Westra, Aviram, and Doell, 2011; Westra, Constantino, and Antony, 2016), than as a standalone treatment, and combined treatment approaches which MI are advancing (Gates, Sabioni, Copeland, Le Foll, and Gowing, 2016; Hogue, Henderson, Becker, and Knight, 2018; Marker and Norton, 2018; Riper et al., 2014). Motivational Interviewing is often applied at the beginning of treatment and later combined with Cognitive Behavioral Therapies (CBT) as in the project COMBINE (Anton et al., 2006) and in the study reported on in this article: the Elderly study (Andersen et al., 2015).

How is the fidelity to MI in these combined therapies? The combination between a more directional therapy with MI has raised some concerns for internal therapist conflicts on which behavior prevails in different situations (Aviram, Westra, Constantino, and Antony, 2016; Boyer, MacKay, McLeod, and van der Oord, 2018; Moyers and Houck, 2011). Findings of fidelity to MI by previous versions of the MITI in these combined settings are mixed (Magill et al., 2018; Aviram et al., 2016; Boyer et al., 2018). The risk of even larger variation in fidelity to MI in combined settings is evident and, thus, measuring fidelity in these treatment approaches is important.

1.5. Aim

The aim of this study was to examine associations between the fidelity to MI and client drinking outcomes for Danish elderly treated with Motivational enhancement therapy (MET) and the Community Reinforcement Approach-Senior (CRA—S) for Alcohol use disorders. As the CRA-S was a combined treatment of MI, cognitive behavioral elements, and recreational elements, this study also explored MI-fidelity in this setting. We hypothesized, that higher levels of treatment fidelity, as measured by the MITI 4, would be associated with improved drinking outcomes.

2. Method

2.1. Design

Data stem from The Elderly Study (Andersen et al., 2015; Sogaard Nielsen et al., 2016), an international RCT, investigating two alternative treatments for addressing alcohol use problems in treatment-seeking elderly patients. Both treatments included four sessions of motivational interviewing, one of them expanding to include additional sessions using the Community Reinforcement Approach (Meyers, Roozen, and Smith, 2011) as well as modules focused on aging (Andersen et al., 2015). The study was conducted in 2014–2016 at outpatient sites in Denmark, Germany and the USA.

2.2. Sample

2.2.1. Participants

Inclusion criteria for participants in the Elderly Study were: 60 or more years of age; positive for alcohol use disorder by DSM 5 (American Psychiatric Association, 2013), and able to pass a comprehension quiz where a minimum of eight out of ten questions about the information given about the study had to be correct (Andersen et al., 2015). Exclusion criteria were: current psychotic symptoms; severe depression; manic episode in the last three months prior to screening; dysregulated bipolar disorder; suicidal behavior at the time of screening; use of illicit opioids or illicit stimulants; under guardianship; and participated in alcohol treatment programs within the last 30 days (Andersen et al., 2015).

All sessions from the Danish Elderly Study were audio-recorded, and 20% were randomly drawn and rated with the MITI 4. The Elderly Study was approved by the National Committee on Health Research Ethics in Denmark (ID number: S-20130138).
2.3. Procedure

2.3.1. Intervention

All participants were offered four sessions of motivational enhancement therapy (MET); moreover, participants randomized to the extended treatment, were offered an additional eight sessions of the Community Reinforcement Approach-Senior (MET + CRA-S) (Andersen et al., 2015). The participants in both groups attended one session per week. Both interventions were based on MET (Anton et al., 2006; Miller, 1995) with the extended version (MET + CRA-S) incorporating principles from Cognitive behavior therapy and the Community reinforcement approach (Meyers et al., 2011). Both interventions were conducted in individual face-to-face sessions, lasting 45–60 min.

Motivational Enhancement Therapy is MI combined with feedback. MI is a non-directive personalized counseling style which aims to explore and elicit participants’ own reasons for change and, thereby, increase motivation for change. The relational approach in MI is particularly important and emphasizes a participant’s feeling of acceptance as fundamental for change. Acceptance is provided through accurate empathy, absolute worth, affirmations, and autonomy (Miller and Rollnick, 2012). Therefore, the MI therapist will first engage the participant in treatment, next focus along with the participant on the direction of therapy, then evoke change talk in this direction, and finally plan how to implement change. Feedback in the Elderly Study was from questionnaires answered by the participants during the baseline interview. In MET feedback is provided in an MI-adherent way and used as an instrument to evoke change talk from the participant (Miller, 1995).

The extended condition followed the format of the Combined Behavioral Intervention (CBI) from The Combined Research Project (Anton et al., 2006). The CRA-S consisted of modules from the Combined manual, adjusted to elderly. In addition to the adjusted modules from the Combined manual, an additional module, focusing on coping with aging, was developed and offered as part of the CRA-S condition (Andersen et al., 2015). Five different treatment modules were available for the extended sessions 5–12. The therapist and participant discussed between them which modules to apply and for how many sessions. This way, some participants could have several sessions of the same module and none of the others. At the same time, it was possible for the therapist to apply a maximum of two modules within the same treatment session. A short description of each module, provided in the extended (MET + CRA-S) treatment, is listed in Table 1. Treatment manuals are available at the University of Southern Denmark website. The manuals for session 1–4 (MET-sessions) only differed on a few aspects between the standard and the extended condition. In the standard MET-only condition, session 4 finalized treatment with a self-change-plan, which the participant should continue to implement on his own. In the extended MET + CRA-S manual, the goal of session 4 was a treatment plan which the therapist and participant will continue to work on in the CRA-S sessions (sessions 5 to 12).

MI was described as the underlying treatment approach throughout the course of both MET sessions and CRA-S sessions. In the manual for CRA-S each module is described, and several examples are given as to how to deliver the modules in an MI-adherent way. Examples from the CRA-S-manual are: “Praise any and all steps the patient has taken to learn and apply new skills.”; “Give plenty of positive reinforcement in the practice process - point out what the patient did well.”; “Use the terminology that is comfortable for your patient.”; “... ask your patient how keeping these records for a while might be beneficial, eliciting change talk.”; “Ask for the patient’s own ideas about how to keep more complete records in the week ahead.”; “Use an evocative style to assess which strategy to use. Rely on open questions that allow the patient to reflect on their own experiences”; and “Remember that you will be using an evocative style to elicit these ideas, rather than supplying them to the patient.” Furthermore, the specific examples of conversations between therapist and participant demonstrate MI-adherent behavior with several seeking collaborations, reflections of change talk, affirmations etc.

2.3.2. Training

Seven therapists were trained by a “train the trainers” approach with continuous supervision (Andersen et al., 2015). The training and supervision took place face-to-face in regular meetings (4 h every 4 weeks), but also as specific and regular written feedback on audio recordings of the sessions. The feedback was given both on randomly picked sessions and on request. Requests occurred when therapists wished for specific feedback in relation to a client or a situation. The supervision addressed all aspects of the interventions: proficiency in MI in all sessions, how to deliver feedback in an MI-proficient way in sessions 1–4, how to perform the modules in sessions 5 to 12, and how to deliver the contents of the modules in an MI-proficient way in sessions 5 to 12. At one of the sites, the participant rotated between therapists: all participants at this site had sessions interchangeably with the three therapists. The researchers performing the follow ups, the therapists, the raters, and the participants themselves were not masked regarding the treatment condition assigned to the participants.

Training the raters in the use of the MITI 4, and details of the fidelity measurement are described in Moyers et al. (2016) and in detail for the

Table 1

<table>
<thead>
<tr>
<th>Module name</th>
<th>Focus</th>
<th>Examples of exercises</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood management training</td>
<td>Managing negative emotions</td>
<td>Analysis of situations with negative emotions. Identifying negative automatic thoughts (NAT). Practicing new ways to cope with focus both on avoiding situations that provoke NAT and constructive thoughts and behaviors after NAT have occurred. Identifying triggers of urge and craving. Training strategies like avoiding, escaping, distracting and enduring of situations with craving or urge. Writing a letter to important people. Practicing how to ask for support from significant others. Inviting significant others into treatment. Finding possible meaningful recovery programs in the local area. Exploring different recreational activities. Trying out possible identified activities. Identifying things that are meaningful. Identifying values. Writing letters to grandchildren or lost ones. Identifying toxic thoughts.</td>
</tr>
<tr>
<td>Coping with craving and urges</td>
<td>Coping with craving, urges, and social pressure to drink.</td>
<td></td>
</tr>
<tr>
<td>Building a sober network.</td>
<td>Increasing positive support from others and engaging in recovery programs.</td>
<td></td>
</tr>
<tr>
<td>Social and recreational counseling</td>
<td>Finding pleasant recreational activities which do not involve drinking</td>
<td></td>
</tr>
<tr>
<td>Coping with concerns related to aging</td>
<td>Coping with loss and sadness associated with aging</td>
<td></td>
</tr>
</tbody>
</table>
present study in Kramer Schmidt et al. (2019). To assess inter-rater reliability, 12% of the recordings, selected for review with MITI 4, were rated by all five raters. Thus 20% of the full sample of recordings at the Danish site were rated for fidelity and 12% of these were rated by all raters. Raters were masked as to which recordings were being simultaneously rated by more than one of them.

2.4. Measures

2.4.1. Outcome variables

Alcohol use was measured by the Form 90 (Miller and Del Boca, 1994) – a reliable measure of the daily use of alcohol 90 days before baseline and until follow-up (Tonigan, Miller, and Brown, 1997). The following outcome variables on alcohol use in the 30 days prior to 26 weeks follow-up were calculated: 1) Percent days abstinent (PDA), 2) Drinks per drinking day (DDD), 3) Percent Heavy Days of Drinking (PHD), and 4) Abstain or not (ABS) (defined as no alcohol intake in the last 30 days). One drink was defined as 12g of alcohol. Heavy drinking was defined as four or more drinks per day for women and five or more drinks per day for men. Finally, a psychosocial outcome was included: 5) Results from The Drinkers Inventory of Consequences (DrnC) (Forcehimes, Tonigan, Miller, Kenna, and Baer, 2007).

2.4.2. Predictor variables – fidelity to MI

MI-fidelity was measured with the Motivational Interviewing Treatment Integrity manual version 4.2.1 (MITI 4) (Moyers et al., 2014). The MITI 4 measures 10 different behavioral counts and 4 global scores through a randomly selected 20-minute segment of the audio-taped session. The behavior counts are: giving information, question, simple reflection, complex reflection, affirmation, seeking collaboration, emphasize autonomy, persuade with permission, persuade, and confront (Moyers et al., 2014). The global scores are evaluated on a five-point Likert scale, while discrete behaviors are counted. The four global scores are: cultivating change talk, softening sustain talk, partnership, and empathy (Moyers et al., 2014). A short description of the measures of the MITI 4 is provided in Table 3. The constructs in the MITI 4 have demonstrated acceptable interrater reliability (Moyers et al., 2016) for both global and behavior counts. Four variables derived from the MITI 4 codes were each used as continuous predictor variables: the average of the global scores empathy and partnership (Relational score); the average of the global scores cultivating change talk and softening sustain talk (Technical score); percent complex reflections of all reflections; and reflection to question ratio. These variables have expert benchmarks recommendations in the summary scores of the MITI 4 for fair and good MI (see Table 4 in the result section). Moreover, two dichotomous variables were created from the benchmark recommendations: a group fulfilling all four benchmarks for fair MI and a group fulfilling all four benchmarks for good MI. Naturally, these two dichotomous variables were not independent; hence, they were analyzed separately. In addition, all MI Adherent behaviors (MIA: affirmation, seeking collaboration and emphasize autonomy), MI Non-Adherent behaviors (MINA: persuade and confront), and the neutral code persuade with permission, were used as predictors for session outcomes.

2.5. Statistics

Mixed effects linear regressions were used to analyze associations between fidelity predictors and the five treatment outcomes. This analysis considered that the participants were nested within therapists and accounted for the baseline level of drinking for each participant. At one site, the therapists were rotating; therefore, the three therapists there had to be clustered together as one. To have an estimate of PDA, DDD, and PHD at baseline, not affected by the treatment seeking itself, they were calculated from use of alcohol from day 31 to day 60 before the baseline interview (time for treatment start). The analyses were performed unadjusted, adjusted for age and gender, and adjusted for all the potentially modifying variables. Potentially confounding variables accounted for: age, gender, treatment condition (MET or MET+CRA), time in treatment, and baseline value of the Alcohol Dependence Scale. Analyses were Bonferroni adjusted, yielding a p-value of 0.001 for significance (0.05/50). Finally, correlations between predictors were estimated and to avoid interactions between the predictors, each predictor was analyzed separately.

The 20% sessions, which were randomly drawn from the full sample of sessions to be rated for fidelity, yielded several sessions on some participants. In the cases where more than one session was rated for a participant, only one session was randomly selected to enter the analysis. In this way, independency of data was secured. The number of participants lost to follow-up was 23 and 47 at 12 weeks and 26-weeks follow-up after baseline, respectively. Because of the higher adherence at 12-weeks follow up, the analysis was also performed with outcome measured at this time point. To check whether the missing data were missing at random, baseline demographics were compared between the two groups. To take missing data into account, an analysis where the missing data is collected from the last meeting point with the participant, i.e. last observation carried forward, was considered. One of the assumptions of a last observation carried forward - analysis is that it must be a conservative estimation of outcome. The effects of MI are usually higher just after treatment; hence, observations carried forward from the 4 and 12-weeks follow-up would not be conservative in these outcomes. Finally, chi-square test, Fisher's exact test, t-tests and two sample Wilcoxon rank sum tests were used, when relevant.

As recommended, interrater reliability in the coding team was assessed by calculating a (1) two-way mixed effects model, (2) consistency of agreement, and (3) average measures intra class correlations coefficient (ICC) for each of the MITI 4 variables (Hallgren, 2012; Jelsma, Mertens, Forsberg, and Forsberg, 2015). To visualize the degree of measurement error, ICC on both the transformed as well as the raw scores were calculated (Hallgren, 2012). Level of interrater reliability was determined using Cicchetti's cut scores (Cicchetti, 1981).

3. Results

3.1. Samples

All in all, the 341 participants and 7 therapists at the Danish sites had a total of 2127 sessions; of which 423 were randomly drawn and coded with the MITI 4. Within these 423 coded sessions, 238 participants were represented. Demographics of these 238 participants are listed in Table 2. At baseline level, the participants who had missing data at 26 weeks, had a higher tendency to be not cohabiting – 68% as compared to 51% (p = 0.04), and to drink more drinks per drinking day (p = 0.03) with a mean of 12.38 (95%CI, 9.86, 14.92) drinks per drinking day as compared to 9.04 (95%CI, 8.13, 9.96). No other demographic differences were found between the groups.

Of the 423 sessions, 52 (12%) were coded by all 5 raters. The ICC-scores were on average excellent (0.78), with a range from 0.50 to 0.96 on the fidelity measures (see Table 3). To illustrate variability in the predictors, median and quartiles are also presented in Table 1. Emphasis autonomy and confront were excluded from further analysis, due to low levels of interrater reliability.

Pearson’s correlations between all MITI 4 measures are presented in the correlation’s matrix in the Appendix A. Listing some of these with Pearson’s r in brackets: Persuade and confront (r = 0.35), persuade and empathy (r = -0.43), persuade and partnership (r = -0.61), complex reflections (CR) and simple reflections (r = 0.48), CR and cultivating change talk (CCT) (r = 0.42), CR and empathy (r = 0.44), partnership and empathy (0.60), seeking collaboration and partnership (r = 0.37), CCT and softening sustain talk (r = 0.36), CCT and partnership (r = 0.56), CCT and empathy (r = 0.53), CCT and the relational score (r = 0.61). The in general low and some moderate levels of correlations indicate that the MITI 4 measures are distinct and at the same time the direction of the
3.2. Treatment fidelity

The mean values of all MI summary scores were above or close to the recommended benchmark values for good MI (Table 4) (Moyers et al., 2014). Of note, 27% of the sessions rated met all four of the recommended benchmark values for good MI in the summary scores, 72% of the sessions met all four of the recommended benchmarks for fair MI, and only 3% met none of the four recommended benchmarks for good MI. Furthermore, the global ratings were significantly better in more MET sessions than CRA-S sessions (also displayed in Table 4). The combined MIA's and MINA's had a mean (95%CI) of 2.19 (1.98; 2.40) and 0.74 (0.60, 0.89) across all rated sessions. The level of MIA's were almost the same across MET (mean: 2.27, 95%CI: 1.99; 2.54) and CRA-S (mean: 2.61, 95%CI: 1.71; 3.21), sessions, but the level of MIA's were somewhat lower in MET sessions (mean: 0.61, 95%CI: 0.46; 0.76) as in CRA-S sessions (mean: 0.96, 95%CI: 0.68; 1.23). There were no significant differences between mean levels of MIA's and MINA's in MET-sessions compared to CRA-S-sessions. Exploring whether MIA or MINA's were present in session at all, 78% of the rated sessions presented at least one MIA in the MET sessions and this number was 77% in the CRA-sessions; whereas, MINA's were present in 30% and 42% of the MET and CRA-S sessions, respectively. The difference in number of sessions rated with at least one MINA was significant (p = 0.01) tested by chi-squared test.

3.2.1. Treatment fidelity and effect of treatment

The assumption of normally distributed residuals was not fulfilled; consequently, we applied a bootstrapped model. Contrary to our hypotheses, no significant associations were found between fidelity measures of MI and drinking outcomes at 26 weeks after treatment (Table 5). The analysis of outcome at 12 weeks follow-up also did not demonstrate significant associations.

3.3. Exploratory sub analyses

Since we found significant differences in the fidelity to MI in the MET-sessions and CRA-S- sessions, we performed the same, but...
separate, analysis within these two groups. The findings for the MET sessions \((n = 165)\), only, were the same as the primary analysis. Surprisingly, and only within the CRA-S condition \((n = 73)\), the binary predictor of whether the session fulfilled all four recommended benchmark values for good MI, was significantly associated with drinking outcome at 26 weeks follow-up. The participants who received therapy that fulfilled all four benchmark values for good MI in the CRA-S-condition, had significantly worse drinking outcomes, with a drop in PDA from baseline to 26 weeks of 40%-points \((95\% CI: -66; -14)\), \((p = 0.003)\), and an increase in PHD of 50%-points \((95\% CI: 22; 77)\), \((p = 0.001)\). This finding was not evident when the analysis was repeated for 12-week drinking outcomes.

4. Discussion

In the present study, the fidelity to MI elements in treatment was carefully and reliably measured in sessions from 238 participants in a randomized trial exploring treatments for Alcohol use disorder. Contrary to our hypotheses, no associations were found between elements of MI and outcomes of drinking 26-weeks after baseline despite high levels of therapist fidelity. Hence, long term predictive validity of the MITI 4 in a setting with treatment seeking Danish +60-year-olds suffering from alcohol use disorder by DSM5, could not be established.

4.1. Fidelity to MI

Shortly, to reiterate the benchmarks for fair and good MI according to the MITI 4, respectively: 3.5 and 4 on the relational summary score, 3 and 4 on the technical summary score, 40%, and 50% complex reflections, and, finally, the reflection to question ratio is recommended to be at least 1 for fair MI and 2 for good MI. Overall, the median levels of each of the summary score in this study fulfilled the proposed benchmarks for good MI. Additionally, 80% of the rated sessions fulfilled the benchmark for good MI on the relational summary score, 67% for the technical summary score, 76% of the rated sessions fulfilled benchmarks for the summary score of percent complex reflections, and 47% of rated session the reflection to question ratio. All in all, these indicate high levels of fidelity and are comparable to other studies with high levels of fidelity to MI \((Roman et al., 2018; Freira et al., 2018; Magill, Graves, et al., 2018; Owens et al., 2017; Serrano et al., 2017)\) and substantially higher than studies with expected lower fidelity to MI \((Copeland et al., 2018; Simper, Breckon, and Kilner, 2017)\). The ratings indicated use of MIA’s, but the general low ratings of emphasize autonomy questions the use of this important MI-technique in this study. Mean levels close to zero of emphasize autonomy is usual in trials with trained, but not experienced therapists \((Theresa B. Moyer et al., 2016; Serrano et al., 2017)\). It could be that this technique indicates more experienced therapists. To report the percentage of sessions fulfilling all four benchmarks for good MI is new, but it is our estimate that 27% of sessions fulfilling all four benchmarks is also in the high range. In support of this, 72% of sessions rated in our study met all four benchmarks for fair MI – the level expected for therapists trained, but not experienced in MI. \((Moyer et al., 2014)\). Hence, these benchmarks of the summary scores reflect the therapists’ level of experience in this study well.

4.2. Mechanisms of MI

One explanation for the lack of associations is that the hypothesized mechanisms of these elements of MI have not been correctly identified. In other words, MI is effective for reasons other than those that have been prescribed \((Miller and Rollnick, 2014; Miller and Rose, 2009)\). Alternatively, it is possible that we are not measuring the elements appropriately. The technical element of MI has been supported \((Copeland, McNamara, Kelson, and Simpson, 2015; Magill, Apodaca, et al., 2018; Pace et al., 2017; Romano and Peters, 2014)\) as an indirect

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Treatment fidelity to MI measured by MITI 4 at the Danish site of the Elderly Study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary scores from the MITI 4</td>
<td>Recommended value for good MI</td>
</tr>
<tr>
<td>Relational: average of empathy and partnership</td>
<td>4 (4.0-4.5)</td>
</tr>
<tr>
<td>Technical: average of cultivating change talk and softening sustain talk</td>
<td>4 (4.0-4.5)</td>
</tr>
<tr>
<td>Complex reflections of all reflections</td>
<td>0.59 (0.50-0.69)</td>
</tr>
<tr>
<td>Reflection to question ratio</td>
<td>1.81 (1.26-2.58)</td>
</tr>
<tr>
<td>Difference between the sessions fulfilling the benchmark criteria for good MI between MET and CRA session is significant: (p &lt; 0.05) and (p &lt; 0.001).</td>
<td></td>
</tr>
</tbody>
</table>
pathway whereby MI consistent behavior on the part of the therapist increases change talk, predicting better client outcomes. However, MI consistent behaviors on the part of the therapist typically increase change talk, predicting better client outcomes. However, MI consistent behavior on the part of the therapist assists and evokes this important transition in client language across the session - from high levels of sustain talk and low levels of change talk in the beginning of the sessions to low levels of sustain talk and high levels of change talk in the end of sessions (Miller and Rollnick, 2012). As proposed by the authors in an earlier paper (Kramer Schmidt et al., 2019), a tallying of therapist behavior increasing sustain talk, like reflections of sustain talk which have been associated with worse outcomes (Lindqvist, Forsberg, Enebrink, Andersson, and Rosen Dahl, 2017) could provide additional information on this subject. In our study, the technical score was significantly higher in the MI sessions with feedback (MET) than in the combined setting (CRA-S); thus, the technical score does seem to imply MI-techniques, which was expected to be performed more frequently in MET. But further validation of the measures of the MITI 4 is necessary.

4.3. Low ambivalence and MI

Additionally, it is possible that the lack of an observed association between treatment fidelity and drinking outcomes could be due to the nature of the sample. Typical participants at the Danish site of the Elderly Study had already made large changes in their drinking prior to the onset of the treatment, a finding not uncommon in clinical trials for people with alcohol use disorder (Anton et al., 2006; Project MATCH Research Group, 1997a). Given their low ambivalence about changing, it is possible that MI sessions in the Danish drinkers should have focused more attention on planning, rather than addressing the ambivalence that is more common in real-world clients who are considering a change in their drinking. In fact, there are studies indicating a discordant effect of MI for individuals who are not ambivalent, but rather committed to the change they have already initiated (Kahler et al., 2004; Kuerbis, Houser, Levak, Shao, and Morgenstern, 2018; Project MATCH Resear ch Group, 1997b; Rohsenow et al., 2004).

If the associations between fidelity and outcomes are moderated by client ambivalence, future analysis may take this into account. What is the use of the MITI 4 in this setting? The importance of measuring and reporting fidelity to MI remains the same. If the global softening sustain talk were to be rated as a default in sessions with no client sustain talk instead of a 4 on the global scale (Kramer Schmidt et al., 2019), this would provide information on client sustain talk and, thereby, an estimate of client ambivalence. This could be an alternative to applying a time-consuming instrument as the MIC (the Motivational Interviewing Skills Code). Opposing this suggestion is the fact that the softening sustain talk measure has not been found associated with client sustain talk (Moyers et al., 2016), but this could be due to low levels of sustain talk in the study by Moyers et al. (2016) as also seems to be problematic in the present study. A rethinking of this measure all in all may be appropriate.

4.4. Softening sustain talk

The revision of the MITI 4, which yielded the softening sustain talk scale, is relatively new (Moyers et al., 2014) and the softening sustain talk scale is also the least reliable of the measures that comprise it (Copeland et al., 2018; Kramer Schmidt et al., 2019; Moyers et al., 2016). This is particularly apparent in the confusion that occurs with this scale when sustain talk is not apparent in a session. Clients who are high in both change and sustain talk (being ambivalent) are expected to be common in MI sessions. Similarly, clients who are high in sustain talk would not be unusual referrals to a motivational treatment. But, what about clients who are no longer ambivalent and are offering little sustain talk, as is the case in the populations of this study? In this case, the rater is attempting to measure the therapist’s impact on a behavior that did not happen. This requires a much more complicated series of decisions about the score that should be given, a process unlikely to be reliable and certainly not standardized within the MITI 4. A revision of the softening sustain talk is needed, to remove the confusion that occurs when clients offer little or no sustain talk in MI sessions. Because sustain talk is an increasing focus of attention to researchers interested in mechanisms of action for MI, it will be important to create measures of therapist competence in responding to sustain talk independent of the client’s baseline offering of this type of language.

4.5. MI fidelity in the CRA-S

In spite of generally high levels of MI also in the combined condition, combining MI with other treatment approaches may compromise elements of MI as indicated by the lower fidelity to global measures in the CRA-S sessions (Table 4). Additionally, significantly fewer sessions contained MINA’s in the MET as the CRA-S sessions. As the absence of MINA’s has been found even more important for effectiveness of MI, this
is an important finding (Gaume, Gmel, Faouzi, and Daeppen, 2009). A lower fidelity to the global measures alone or higher level of MINA’s was not found associated with treatment outcome in this study, hence, this lower fidelity to MI in the CRA-S-sessions was not associated with worse drinking outcomes. The MITI 4 was designed to be more MI-specific; thus, the global measures and the absence of MINA’s seems to indicate a purer MI-condition as provided in the MET. This adds to the validity of the instrument.

In different settings, some studies of treatment with MI combined with Cognitive Behavioral Therapy (CBT) report fair levels of fidelity to MI, based on previous fidelity measures, but finding are mixed (Barrera, Smith, and Norton, 2016; Boyer et al., 2018; Magill, Graves, et al., 2018; Westra et al., 2016) and one study report the fidelity to CBT compromised (Westra et al., 2016). Combining MI with more structured and expert-driven treatments may present dilemmas that necessarily compromise the conditions of autonomy-support, that define MI (Moyers and Houck, 2011). The higher tendency in our study to persuade and confront in the combined setting supports this. Combining MI with less structured treatment may, on the other hand, increase the impact of MI (Boyer et al., 2018). In particular, difficulties may arise concerning what the therapist ought to do when the client is not accepting or following the treatment plan, not an uncommon occurrence in interventions for alcohol use disorders (Moyers and Houck, 2011). In the Elderly study, there was strong emphasis on MI, especially in session 1–4, and if the participant did not accept a treatment plan this was faced in an MI adherent way by accepting and postponing the decision for later. In the description of the modules where MI is combined with cognitive behavioral and community reinforcement approaches, the emphasis on MI varies. It is mentioned more sporadically, and some modules are described in a way that does not seem to leave a choice for the therapist or the participant in the treatment progress. The therapist is more often the expert and has a lot of questions. It does not seem possible to back down from a task of making the client try out a recreational activity or another behavior in order not to induce negative automatic thoughts. The language is more imperative both in the description as well as in the examples of how to talk with the client. The modules vary a lot, some of them remember more of MI in their descriptions than others. One module, where notions on how to incorporate MI seems far from each other, is the Mood module—one of the most employed in the extended sessions.

A study by Aviram et al. (2016) supports the importance of timing in adherence to MI. In their study of MI combined with CBT for anxiety disorders, MI adherent behaviors during times of client resistance in sessions were associated with outcome, whereas high levels of MI adherent behaviors in general were not (Aviram et al., 2016). Therefore, the combination of MI with other therapies may require even more expertise and sensitivity, on behalf of the therapist, to cues from the client on when to use which approach (Aviram et al., 2016).

This may also explain our finding, where higher MI fidelity in sessions of the Community Reinforcement Approach Senior (CRA—S) was associated with more client drinking. It is a tentative finding, though, because it was an explorative analysis, in which one of the results had a p-value higher than that allowed by the Bonferroni correction, and this result was not found at 12 weeks follow up when it should have been more likely to occur. Thus, this result should be taken with caution, as the possibility of a Type I error is not trivial. Speculations about this result could either be because the therapist provided MI even though the client was committed to change and therefore not timing the delivery of high adherent MI behavior correct. This is in line with previous findings of lower effectiveness of MI when the client is already committed to change (Kahler et al., 2004; Kuerbis et al., 2018; Project MATCH Research Group, 1997b). Another explanation could be, that the therapist provided MI because the client was still ambivalent after four sessions of Motivational Enhancement Therapy (MET) and this unresolved ambivalence could be associated with higher drinking level at 26 weeks follow-up. Importantly, as we do not have any measures of fidelity to the other elements of the CRA-S intervention, we cannot rule out an influence from these unmeasured elements on these findings. As an example, the high level of MI could mean lower delivery of cognitive behavioral and recreational approaches, and that this is what worsens outcomes (Constantino, Romano, Coyne, Westra, and Antony, 2018).

4.6. Limitations and strengths

The results of our study carry several limitations with them. First, drinking outcomes in this study were gathered 6 months after treatment initiation – well beyond the time that a brief treatment focusing on motivation would be expected to have its greatest impact. Because it influences the client’s motivation to make a behavior change, the impact of MI should be greatest during and just after treatment (Smedslund et al., 2011). Using a more distal outcome conflates the impact of MI with many other post-treatment events (such as the loss of relationships, economic events and health concerns) that can be important in influencing problem drinking. Thus, the 6-month follow-up point yields a conservative estimate of effectiveness for a treatment intended to influence the client’s immediate motivation to change.

The participants lost to follow up at 26 weeks had more drinks per drinking day at baseline and were less cohabiting. Thus, there is a risk of bias when participants with higher likelihood of positive outcome were assessed more often than the remaining. Paradoxically, the Elderly Study may have been underpowered to detect the differences in outcomes that were present because of the high quality of the intervention, as it was delivered by study therapists. Treatment effects for behavioral interventions are, typically, very small (Wampold, 2015), so it is likely that the fidelity to the treatment may represent an even smaller effect which it would require a large sample of therapists and participants to detect. With 238 clients and seven therapists, it is entirely possible that very small differences, especially in the absence of variability in treatment delivery among therapists, would lead to a Type II error. That is, we might have been unable to observe an effect in our small sample even if it truly exists in the population. Moreover, the possible conflicting associations between MI-fidelity and outcome in the CRA-S sessions may camouflage associations which we were unable to detect in the even smaller sample of only MET-sessions (n = 165).

Our finding of high levels of fidelity to MI and low variability in a manualized settings with high levels of supervision is not surprising (Hallgren et al., 2018). This, and other efficacy elements like frequent meetings between MITI 4-raters and high levels of organizational support in treatment institutions, limits generalizability to real world settings.

Another drawback of this study is that we do not have any measures of the adherence to the other elements of the CRA-S-condition and how well the therapists integrated the different therapeutic elements of this condition. This would provide important insight on how therapists should proceed in situations that generate opposing guidelines within combinations of MI and other therapies. Additionally, this information could shed light on the reversed findings in the CRA-S sessions: the timing of MI in this combined setting and the influence on outcome of the other elements in the CRA-S would be especially interesting to unravel. An influence that is likely to have affected our findings.

Finally, these findings are from a fidelity measurement in a Danish speaking setting with MI delivered to participants above 60 years of age, only. It is natural and assumed that MI therapists adapts MI to the norms and values of the population it is applied (W.R Miller & S Rollnick, 2012). We do not know how the fidelity to MI changes according to different languages. Several studies have been performed outside North America and in languages other than English with previous versions of the MITI (Beckman et al., 2017; Brueck et al., 2009; Eno Persson et al., 2016; Jamieson, Bradshaw, Lawrence, Broughton, and Venner, 2016; Kouwenhoven-Pasmooy et al., 2018; Lee, Tavares, Popat-Jain, and Naib, 2015; Lindqvist et al., 2017; McCambridge, Day, Thomas, and Strang, 2011; Mesters, Keulen, de Vries, and Brug, 2017;
5. Conclusion

The MITI 4 measures reliably in almost all measures in a Danish outpatient setting treating +60-year old people for Alcohol use disorder. The fidelity to MI in the interventions delivered at the Danish site of the Elderly Study was high and with low variation. Global measures of the MITI 4, and thereby fidelity to MI, was higher in sessions of Motivational Enhancement Therapy than in the combined condition where MI was combined with cognitive and community recreational approaches. The population had low levels of ambivalence, hence, risks that MI did not have the expected impact. Predictive validity of the MITI 4 on alcohol use outcomes was not confirmed. The technical global measures of the MITI 4 is suggested revised to improve measures of evocation competencies from the therapist; especially, the global measure softening sustain talk.

Conflicts of interest

None.

Acknowledgements

Our gratitude for the financial support provided, unconditionally, by the Lundbeck Foundation, The Region of Southern Denmark and The University of Southern Denmark. Thank you to the participants, therapists and institutions for taking part in this study. Thanks to members of the coding team: Birgit Jensen, Peter Gabelgaard Brolund, Louise Bungsdgaard, and Ayse Corap; moreover, thank you to the statisticians in this project: Anna Mejldal, Anders Bo Bojesen, and Christian Sibbersen. Finally, thank you to Annemette Munk Svensson for proofreading this paper.

References


60+ years: Design of a randomized clinical trial conducted in three countries (Elderly Study). BMC Psychiatry, 15(1), 280.


