Published in final edited form as:

J Subst Abuse Treat. 2015 April ; 51: 47–52. doi:10.1016/j.jsat.2014.10.009.

Comparing the Motivational Interviewing integrity in two prevalent models of brief intervention service delivery for primary care settings

Chris Dunn, Ph.D.^a, Doyanne Darnell, Ph.D.^a, Adam Carmel, Ph.D.^a, David C. Atkins, Ph.D.^a, Kristin Bumgardner, B.S.^a, and Peter Roy-Byrne, M.D.^a

^aDepartment of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, USA.

Abstract

This quasi experimental study compared the Motivational Interviewing (MI) integrity in two prevalent brief intervention (BI) service delivery models for drug abuse. Routine primary care providers (RCPs) and non-routine care providers (NRCPs) performed BIs using an MI style within the same medical setting, patient population, and Screening, Brief Intervention, and Referral for Treatment (SBIRT) protocol. Interventionists (9 RCPs and 6 NRCPs) underwent similar MI training and performed a total of 423 BIs. We compared the MI integrity scores for all audio recorded sessions from these two SBIRT models for up to 40 months post MI training. Both groups met beginning proficiency in MI on 4 of 5 MI integrity scores, but NRCPs met more of the higher competency criteria than RCPs. There may be limitations with regards to MI fidelity when using RCPs to conduct BIs in some primary care settings. Further experimental investigation is warranted to replicate this finding and identify casual factors of observed differences in MI fidelity.

Keywords

Motivational Interviewing; integrity; Screening Brief Intervention and Referral for Treatment (SBIRT); substance use; primary care

1. Introduction

The Joint Commission on Accreditation of Healthcare Organizations has developed performance measures for alcohol and tobacco screening and may soon require hospitals to provide alcohol and tobacco SBIRT services for all inpatients (Joint Commission on Accreditation of Healthcare Organizations, 2012). However, if such an enormous

© 2014 Elsevier Inc. All rights reserved.

Address correspondence to: Chris Dunn, Ph.D., Harborview Medical Center, Box 359911, 325 Ninth Avenue, Seattle, WA 98104, USA. Tel.: +1 206 251 4890. cdunn@uw.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

implementation were to fail to achieve clinical effectiveness due to poor counseling, the public health benefit that SBIRT was designed to achieve would be lost. Poor counseling during the BI is clearly a threat to SBIRT effectiveness (Miller & Rollnick, 2014). Yet 30 years of SBIRT research reveals little about which individuals or SBIRT implementation models produce the best BI counseling (Nilsen, 2010). Therefore, this study compared the counseling integrity produced by two convenience samples of interventionists: routine care providers (RCPs) with competing job demands in the clinic other than SBIRT (e.g., Desy, Howard, Perhats, & Li, 2010; Seale, Shellenberger, & Clark, 2010) and non-routine care providers (NRCs) with only SBIRT duties (e.g., Gryczynski et al., 2011).

After decades of research testing the effects of using mostly physicians to perform BIs, the physician-based model has failed at the implementation stage for a reason that we focus on in this study: competing clinical priorities resulting in limited time and/or motivation to perform BI counseling (Babor & Higgins-Biddle, 2000; Babor et al., 2007). Non-physician RCPs, such as social workers or nurse practitioners, may alternatively be tasked with performing SBIRT interventions. Although SBIRT conducted by non-physician RCPs appears to be equally as effective as when conducted by physicians (Sullivan, Tetraault, Braithwaite, Turner, & Fiellin, 2011), issues of limited time and competing responsibilities may still impact these SBIRT interventionists. To address the challenges of competing clinic roles and priorities, SBIRT programs may instead decide to utilize specialists whose only role in the clinic is to perform SBIRT interventions (NRCs). In fact, within primary care settings there is some indication that using NRCs results in a higher rate of completion of BIs with patients who screened positive than using RCPs (Babor, Higgins-Biddle, Dauser, Higgins, & Burleson, 2005).

Competing clinic roles and priorities may negatively impact the quality of the BI counseling. In the U.S., most SBIRT implementations are currently using at least some elements of Motivational Interviewing (MI) to conduct the BI counseling (Agerwala & McCance-Katz, 2012). Yet, MI is difficult to learn (Miller & Rollnick, 2013), and the cost of monitoring MI integrity prevents us from knowing how well counselors are actually doing MI in large SBIRT implementation efforts (Atkins, Steyvers, Imel, & Smyth, 2014). MI requires more acquired listening skills and strategic thinking than brief advice or brief negotiation methods used in other SBIRT protocols (Agerwala & McCance-Katz, 2012). NRCs may have more time for and fewer distractions from attentive listening than busy RCPs who may have the next patient already waiting to see them, or who may not have known until the last minute before greeting a patient that the patient had a positive screen and a BI would be added to the main business of that visit. The professional identity of interventionists and their perceived roles could also affect the effort and focus that they devote to counseling (Roche & Freeman, 2004). It may be that NRCs are motivated to focus more of their attention on adhering to an MI counseling style than RCPs whose SBIRT duties are additions to their larger job descriptions.

In the present study, NRCs were research personnel on the larger research project and had significantly less clinical experience than RCPs. Clinical experience might be expected to influence MI integrity. One of the challenges in learning and counseling in MI to fidelity is the common difficulty of suppressing the “righting reflex” (Miller & Rollnick, 2013). This

term refers to the natural desire to help suffering people by reflexively trying to set right their circumstances. Among healthcare clinicians, this urge commonly drives clinical habits such as searching for and assessing problems, problem solving, and giving unsolicited advice. These habits are inconsistent with the spirit of MI which is to be collaborative, evoke the patient's expertise, and respect patient autonomy to choose whether to change. Old habits stemming from the righting reflex may be harder to break for more seasoned clinicians. Consistent with this premise, Miller and Mount (2001) found that clinicians learned new MI skills more readily than they learned to drop old habits. In contrast, a meta-analysis of MI training found similar learning effect sizes for mental health versus other professionals, and for interventionists with more versus fewer years of clinical experience with a given patient population (de Roten, Zimmermann, Ortega, & Despland, 2013). Furthermore, Baer and colleagues (2004) found that neither academic degree nor years of clinical experience was related to reaching MI proficiency after training.

Despite the fact that NRCs and RCPs have both been used in SBIRT implementation, we are not aware of any study that specifically examines how the use of interventionists with these different clinic roles impacts the quality of the BI counseling. Therefore, this quasi-experimental implementation study compared the MI integrity during brief interventions delivered by RCPs and NRCs using the same SBIRT protocol in the same clinical setting. Given that we expect the greater demands on clinician's time and competing clinical demands to negatively impact MI fidelity, we hypothesized that RCPs would have lower MI integrity than NRCs.

2. Method

2.1. SBIRT protocol of larger study

Data for the current study was generated in the intervention arm of a larger randomized trial of SBIRT for patients screening positive for problem drug use during routine medical visits at primary care clinics in an urban medical center in Washington State (see Krupski et al., 2012 for details on study protocols). Eligible patients were male and female adults attending routine primary care visits who reported using an illegal drug or non-prescribed medication at least once in the past 90 days during screening. Exclusion criteria were current or recent substance abuse treatment, severe cognitive impairment, suicidality, or active psychosis. Consented patients were assessed for drug use severity using the 10-item Drug Abuse Screening Test (DAST-10) (Gavin, Ross & Skinner, 1989). Patients were then randomized to a usual care control group or the BI group, who received a BI from the trained interventionist in the clinic during their medical visit. The BI protocol included giving patients screening feedback using their DAST-10 results and discussing its meaning, exploring consequences of whichever drug patients preferred to discuss, discussing their reasons to change, and planning how they might change their drug use if patients were ready and willing to do so. The Referral to Treatment protocol in which patients received a list of community treatment resources was held constant for intervention group and control group patients.

2.2 Participants

Interventionists were also consented as participants for this study which was approved by the University of Washington Institutional Review Board. Per the original study design, eligible interventionists were those designated by their social work supervisor to work for the study. Therefore early in the study, only medical social workers were trained as interventionists who were employed by the primary care clinics (RCPs) and whose daily appointments with patients involved helping them to access public community resources. To increase study enrollment later in the trial, some of the study's research personnel (NRCPs) were trained to perform BIs and given protected work time to perform BIs in the clinics. Two of the 17 total trained interventionists (both RCPs) actually performed no BIs before leaving the study and were excluded from our analyses, leaving a sample of 15 interventionists (9 RCPs and 6 NRCPs).

2.2 MI training and supervision

All interventionists in both groups were trained by the first author to perform the BI protocol using an MI counseling style. The training occurred in 4 cohorts to accommodate continual personnel changes and patient recruitment goals during the trial. Training method, didactics, and skill practice were held constant, but actual doses of these components varied somewhat due to interventionists' work schedules and patient recruitment pressures. Because coaching and feedback enhances workshop training (Miller, Yahne, Moyers, Martinez, & Pirritano, 2004; Moyers et al., 2008; Madson, Loignon, & Lane, 2009), the training package for each cohort consisted of workshop (MI didactics, demonstrations, and practice) followed by up to five optional weekly audio recorded telephone role-plays with standardized patient actors on which they received emailed feedback from the trainer. In terms of training dose, the RCPs received significantly more hours of workshop training ($M = 9.6$, $SD = 1.7$) than NRCPs ($M = 5.9$, $SD = 2.4$), but not significantly more standardized patient practice sessions. Once patient recruitment and BIs began, the trainer reviewed approximately every fifth audio recorded BI for each interventionist and emailed them feedback during the trial. This feedback did not include any MI scoring data, because this was thought to be too expensive to generalize to most real world implementation efforts. Trainer feedback was most often to increase one's reflections, efforts to evoke and respond to patient change talk, and keep the conversation focused on drug topics. Table 1 shows that there were no statistically significant group differences in the number of BIs on which interventionists received trainer feedback during the study. No face-to-face individual supervision occurred during the trial, but monthly 1.5-hour group supervision sessions with the trainer were held during which one or two selected audio recorded BIs were reviewed by the group and relevant MI skills were discussed and practiced. Interventionists from both groups attended approximately 75% of all group supervision sessions and the average monthly attendance at these sessions was 5 interventionists.

2.3 Measures

All BIs in the study were audio recorded and scored by trained coders using the Motivational Interviewing Treatment Integrity (MITI) scales (Moyers, Martin, Manual, Miller, & Ernst, 2010). We operationalized MI integrity using 6 measures from the MITI.

We used the five summary scores calculated from subscales of the MITI and defined in the MITI manual, to which we added the Empathy subscale, because Empathy measures the quality and depth of the counselor's understanding of the client. Two of these 6 scores are overall session Likert ratings of MI Spirit (an average of 3 global Likert scales: Collaboration, Autonomy, and Evocation) and Empathy. Four other summary scores are ratio scores derived from counts of counselor behaviors: percentage of total reflections that are complex (versus simple reflections), percentage of total questions that are open questions (versus closed questions), the ratio of total reflections to total questions made by the interventionist, and percentage of MI Adherent responses (ratio of Adherent responses to the sum of Adherent and Nonadherent responses). These scores were calculated for each interventionist and then used to calculate group means. The MITI manual includes expert-suggested thresholds for Beginning Proficiency (lower standard) and Competency (higher standard) for these MITI scores, and our results are reported relative to these two thresholds.

2.4 MI coder training

Early in the study, the first author trained two research assistants to perform MITI coding. Coder training consisted of weekly meetings for 6 weeks requiring approximately 40 hours of coder time to learn adherence to the MITI scoring manual (Moyers et al., 2010). Before coding real BI sessions, both coders coded a sample of 17 practice tapes and intraclass correlation coefficients (ICC) were calculated to measure inter-rater reliability for the behavioral counts. For the Likert scales, the percentage of agreement was calculated as a reliability check (see Table 2 for inter-rater reliability results). Later in the study, a third coder previously trained by another project was hired; he initially blind coded the same practice sample of 17 sessions for a second reliability check before proceeding to code real BI sessions. As seen in Table 1, inter-rater reliability coefficients for 5 of the MITI summary scores were considered good to excellent (Cicchetti, 1994). The only exception was the percentage of MI adherent responses score which had poor reliability (0.13) and was thus excluded from the analysis.

2.2. Data Analysis

Because there were a relatively small number of interventionists, initial comparisons of RCP and NRCP groups focused on means and 95% confidence intervals of MITI summary scores for each individual interventionist, which were also compared to MITI Beginning Proficiency and Competency thresholds. The two Likert scale summary scores, MI Spirit and Empathy, were very strongly skewed and thus converted to binary variables (5 versus < 5) for use in statistical analyses examining group differences in counseling. Because MITI scores were nested within therapists, multilevel models were used to estimate differences between means for continuous variables and proportions for binary variables. All analyses were done using R v3.0.3 (R Core Development Team, 2014). One consideration in analyses was whether greater experience with BIs might improve performance and potentially confound RCP vs. NRCP comparisons. To examine this possibility, preliminary analyses explored the total number of BIs performed and BI session length by group and whether these variables were related to MITI scores.

3. Results

As seen in Table 1, RCPs were significantly older and had more years of clinical experience than NRCPs. Preliminary analyses of the total number of BI sessions performed by each group were not significant ($t(12) = 0.25, p = .80$), and the number of BI sessions performed was unrelated to MITI scores. NRCP sessions were significantly longer in duration than RCP sessions (see Table 1), but session duration was not related to MITI scores.

As seen in Table 2, mean scores for NRCPs were significantly greater than RCPs for 4 of 5 MITI summary measures, the only null result being the percentage of Open Questions. Both groups met the Beginning Proficiency thresholds on 4 of 5 MITI scores (Percent Complex Reflections, Reflection-to-Question Ratio, Spirit, and Empathy). Both groups were below Beginning Proficiency for Percent Open Questions. Additionally, NRCPs met the higher Competency thresholds for the same 4 MITI measures, whereas RCPs met Competency for 2 of those 4 measures (Spirit and Empathy).

Figure 1 shows graphically the individual-level data that was analyzed for the results in Table 2, and it also illustrates individual counselor variation. Relative to the other measures, Reflection-to-Question Ratio had the most variability, particularly within the NRCP group. The figure also shows that in the case of the Beginning Proficiency criteria, interventions received by patients very seldom dipped below the cutoff, even at the low end of each counselor's variability. The figure also shows that the global scales of Empathy and Spirit were clearly the strongest scores for both groups.

4. Discussion

This study compared the MI integrity of SBIRT interventionists with two different roles in a primary care clinic. One is a member of the staff and clinical team engaged in SBIRT among other clinical tasks (RCPs), and the other is a specialist employed by an outside source with the only purpose in the clinic of providing SBIRT services (NRCPs). Additionally, on account of being research personnel for the larger study, the NRCPs had fewer years of clinical experience than RCPs. Consistent with our hypothesis, based on MITI summary scores, NRCPs performed MI with greater integrity. NRCP interventionists scored higher on 4 of 5 MITI summary scores, including the two global measures of MI Spirit and Empathy and the two behavioral scores, percentage of reflections that were complex and the ratio of reflections to questions. Both groups met Beginning Proficiency on 4 of 5 MITI scores. However, the NRCP group met the higher Competency threshold on 4 of these variables, whereas the RCP group met only 2 of the 5 Competency criteria.

Our data do not support our contention that greater MI fidelity in the NRCP group is due to NRCPs having more time available to engage patients in MI-style counseling. Although we did observe that NRCP BI sessions were of a statistically significant greater length, BI session length was not related to MI fidelity. A possible explanation for the better performance by NRCPs could be that they were not encumbered by other roles or tasks in the clinics that could have distracted them from the BI counseling (Babor et al., 2005). RCPs' routine patient schedules were often interrupted by impromptu demands whenever a

patient screened positive and was randomized to receive a BI before leaving the clinic. In contrast, NRCPs who were all research staff for the study did not have a preexisting caseload of patients, did not get scheduled for appointments outside of BIs, and usually did not have to rush back to their research duties after a BI because the study coordinator was able to protect some of their time. It would have improved the study to perform exit interviews of interventionists to assess their perceptions of interruptions or competing job demands. Another possible explanation for observed differences in MI integrity could be that the NRCP group was better in MI at baseline. Although the study did not measure MI integrity at baseline, this also seems unlikely, because NRCPs had much less overall clinical experience than RCPs and more formal clinical training has been found to predict better baseline MITI scores (Baer et al., 2009).

Despite that prior clinical experience has not been shown to predict MI fidelity in other studies, prior clinical experience was a major difference between the NRCP and RCP groups in our study and we cannot rule out the possibility of its influence in our findings. In fact, there is some indication that NRCPs had less difficulty learning MI than RCPs, which may be a result of having markedly less previous clinical experience. Specifically, the righting reflex can exacerbate two common counseling habits measured by the MITI, which may be harder to unlearn for more seasoned clinicians: asking too many closed questions and making too few reflections, both of which would decrease the reflection/question ratio. In this study, there was a large group difference after training in the Reflection-to-Question Ratio, suggesting that old habits may be harder to break for counselors with more experience.

Because the MI training occurred in 4 separate cohorts over 2 years we cannot rule out unequal training as an explanation for our findings. For example, most of the RCPs were trained earlier in this period and the NRCPs were trained later, by which time the training method could have had a greater impact due to practice effects. However, this seems unlikely because the only significant group difference in training dose (workshop hours) was in favor of the RCPs. Additionally, we are confident that the differences were not due to differences in the SBIRT protocol, clinic setting, or population served, since these variables were the same in both groups. The study's sample and the characteristics of our two groups limit our ability to generalize to other populations of SBIRT interventionists. Because the RCPs were all medical social workers our results tell us little about how well other primary care clinicians such as psychologists or nurses might perform SBIRT counseling. Similarly, the NRCPs were all research staff, which may not generalize to other primary care NRCP populations. However, the known distinguishing characteristics between NRCPs and RCPs in this study, including differences in prior clinical experience and the fact that one group was employed by the clinic and one employed by the research grant outside the clinic, point to relevant foci of future research on SBIRT implementation in primary care clinics. In particular, our findings suggest careful examination of whether SBIRT interventionists could provide a higher quality BI if their clinic role is defined solely around providing SBIRT services (as was the case for the NRCPs) rather than SBIRT being an additional task added to an already full set of clinical responsibilities (as was the case for the RCPs) is warranted. In terms of generalizability, a final consideration is that MI is not the only counseling style

used for SBIRT in primary care, and our findings may not generalize to other counseling methods used.

Another consideration for this study is that we used a posttest only design without time series measurements at fixed follow-up intervals; therefore, it is unknown how the groups compared before training, immediately after training, or at any fixed time point over the course of the 40-month enrollment period. The groups may well have undergone changes in their integrity rankings during this long period, changes rendered undetectable by our design. Despite not knowing what MI integrity fluctuations may have occurred between the groups over time, we can say that across the life of this study, both groups of interventionists scored high enough often enough for their overall means to meet Beginning Proficiency on 4 of 5 MITI summary measures and seldom dipped below this level over this relatively long period of time.

Despite noted limitations, this study benefits from a number of unique strengths. First, we used of an average of 25 BI sessions per counselor completed over the course of 40 months. This is the greatest number of sessions per counselor coded for MI integrity and the longest period of time post-training reported thus far in the literature, which allowed us to compute a reliable estimate of counselor MI integrity (Dennhag, Gibbons, Barber, Gallop, & Crits-Christoph, 2012). Previous research indicates that counselor skill decreases in the months post-training (e.g., Moyers et al., 2008) and that there is very large within-counselor variation in MI, which may be due to provider, patient, and dyad factors (Imel, Baer, Martino, Ball, & Carroll, 2011; Forsberg, Forsberg, Lindqvist, & Helgason, 2010). Given this variability, counselor integrity is best estimated with multiple work samples (Dennhag et al., 2012), yet this is a rare practice for randomized trials or implementation efforts. Additionally, this study reported on a greater number of the various MITI summary scores than is typical in the literature (5 versus the average reported number of 2.3; Schwalbe, Oh, & Zweben, 2014), providing a more thorough description of MI fidelity.

It is impossible to know if or how our study findings translate to the clinical effectiveness of BIs. There are few studies examining the relation between MITI scores and outcomes, and no individual MITI summary score is yet known to strongly account for behavior change (Bertholet, Palfai, Gaume, Daeppen, & Saitz, 2013). It is also unknown whether the expert-derived MITI Beginning Proficiency and Competency cut-offs themselves translate to different magnitudes of behavior change. Until the day comes when many MI process studies have been funded and carried out, we will never know what makes MI as effective as previous research has shown and what the implications are of any observed differences in MI integrity. Despite limited knowledge regarding the clinical significance of observed MITI score differences, we can say that the consistently lower fidelity observed among RCPs points to a potential limitation of relying on the RCP model of SBIRT delivery. Future research comparing RCPs to NRCPs in real-world SBIRT implementation that is designed to identify causes of observed differences (e.g., groups that vary only by the absence/presence of non-SBIRT clinical demands) is warranted.

Acknowledgments

The research was supported by grants R01 DA026014-01 from the National Institute on Drug Abuse and R01 MH073613 and T32 MH082709 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors wish to thank Jessica Lowe, B.A., Katy Thysell, B.S., and James Darin, B.S., Research Study Coordinators at University of Washington, for their assistance in coding and data management.

References

- Agerwala SM, McCance-Katz EF. Integrating screening, brief intervention, and referral to treatment (SBIRT) into clinical practice settings: A brief review. *Journal of Psychoactive Drugs*. 2012; 44(4): 307–317. [PubMed: 23210379]
- Atkins DC, Steyvers M, Imel ZE, Smith P. Scaling up the evaluation of psychotherapy: Evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science*. 2014; 9:49
- Babor TF, Higgins-Biddle JC. Alcohol screening and brief intervention: Dissemination strategies for medical practice and public health. *Addiction*. 2000; 95(5):677–686. [PubMed: 10885042]
- Babor TF, Higgins-Biddle JC, Dauser D, Higgins P, Burleson JA. Alcohol screening and brief intervention in primary care settings: Implementation models and predictors. *Journal of Studies on Alcohol*. 2005; 66:361–368. [PubMed: 16047525]
- Babor TF, McRee BG, Kassebaum PA, Grimaldi PL, Ahmed K, Bray J. Screening, brief intervention, and referral to treatment (SBIRT): Toward a public health approach to the management of substance abuse. *Substance Abuse*. 2007; 28(3):7–30. [PubMed: 18077300]
- Baer JS, Rosengren DB, Dunn CW, Wells EA, Ogle RL, Hartzler B. An evaluation of workshop training in motivational interviewing for addiction and mental health clinicians. *Drug and Alcohol Dependence*. 2004; 73:99–106. [PubMed: 14687964]
- Baer JS, Wells EA, Rosengren DB, Hartzler B, Beadnell B, Dunn C. Agency context and tailored training in technology transfer: A pilot evaluation of motivational interviewing training for community counselors. *Journal of Substance Abuse Treatment*. 2009; 37:191–202. [PubMed: 19339139]
- Bertholet N, Palfai T, Gaume J, Daeppen JB, Saitz R. Do brief alcohol motivational interventions work like we think they do? *Alcoholism: Clinical and Experimental Research*. 2013; 38(3):853–859.
- Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6(4):284–290.
- Dennhag I, Gibbons MB, Barber JP, Gallop R, Crits-Christoph P. How many treatment sessions and patients are needed to create a stable score of adherence and competence in the treatment of cocaine dependence? *Psychotherapy Research*. 2012; 22(4):475–488. [PubMed: 22449079]
- de Roten Y, Zimmermann G, Ortega D, Despland J-N. Meta-analysis of the effects of MI training on clinicians' behavior. *Journal of Substance Abuse Treatment*. 2013; 44:155–162. [PubMed: 23537923]
- Desy PM, Howard PK, Perhats C, Li S. Alcohol screening, brief intervention, and referral to treatment conducted by emergency nurses: An impact evaluation. *Journal of Emergency Nursing*. 2010; 36(6):538–545. [PubMed: 21078465]
- Forsberg L, Forsberg LG, Lindqvist H, Helgason AR. Clinician acquisition and retention of motivational interviewing skills: a two-and-a-half-year exploratory study. *Substance Abuse Treatment, Prevention, and Policy*. 2010; 5(8):1–14.
- Forsberg L, Ernst D, Farbring CA. Learning motivational interviewing in a real-life setting: A randomised controlled trial in the Swedish Prison Service. *Criminal Behaviour and Mental Health*. 2011; 21(3):177–188. [PubMed: 21706526]
- Gavin DR, Ross HE, Skinner HA. Diagnostic validity of the Drug Abuse Screening Test in the assessment of DSM-III drug disorders. *British Journal of Addiction*. 1989; 84(3):301–307. [PubMed: 2650770]

- Gryczynski J, Mitchell SG, Peterson TR, Gonzales A, Moseley A, Schwartz RP. The relationship between services delivered and substance use outcomes in New Mexico's screening, brief intervention, referral for treatment (SBIRT) initiative. *Drug and Alcohol Dependence*. 2011; 118(2–3):152–157. [PubMed: 21482039]
- Imel ZE, Baer JS, Martino S, Ball SA, Carroll KM. Mutual influence in therapist competence and adherence to motivational enhancement therapy. *Drug and Alcohol Dependence*. 2011; 115:229–236. [PubMed: 21227601]
- Joint Commission on Accreditation of Healthcare Organizations. Performance measurement initiatives: Screening and treating tobacco and alcohol use. 2012 http://www.jointcommission.org/substance_use/.
- Krupski A, Joesch JM, Dunn C, Donovan DM, Bumgardner K, Peregrine Lord S, Ries R, Roy-Byrne PP. Testing the effects of brief intervention in primary care for problem drug use in a randomized controlled trial: Rationale, design, and methods. *Addiction Science & Clinical Practice*. 2012; 7(1):27. [PubMed: 23237456]
- Madson MB, Loignon AC, Lane C. Training in motivational interviewing: A systematic review. *Journal of Substance Abuse treatment*. 2009; 36:101–109. [PubMed: 18657936]
- Miller WR, Mount KA. A small study of training in motivational interviewing: Does one workshop change clinician and client behavior? *Behavioural and Cognitive Psychotherapy*. 2001; 29:457–471.
- Miller, WR.; Rollnick, S. *Motivational interviewing: Helping people change*. 3rd ed.. New York: Guilford Press; 2013. ISBN 1-57230-563-0.
- Miller WR, Rollnick S. The effectiveness and ineffectiveness of complex behavioral interventions: Impact of treatment quality. *Contemporary Clinical Trials*. 2014; 37:234–241. [PubMed: 24469237]
- Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*. 2004; 72:1050–1062. [PubMed: 15612851]
- Moyers T, Manual J, Wilson P, Hendrickson S, Talcott W, Durand P. A randomized trial investigating training in motivational interviewing for behavioral health providers. *Behavioural and Cognitive Psychotherapy*. 2008; 36:149–162.
- Moyers TB, Martin T, Manual JK, Miller WR, Ernst D. Revised global scales: Motivational Interviewing Treatment Integrity 3.1 (MITI 3.1). 2010 Retrieved from: http://casaa.unm.edu/download/MITI3_1.pdf.
- Nilsen P. Brief alcohol intervention—where to from here? Challenges remain for research and practice. *Addiction*. 2010; 105:954–959. [PubMed: 20121717]
- O'Donnell A, Anderson P, Newbury-Birch D, Schulte B, Schmidt C, Reimer J, Kaner E. The impact of brief alcohol interventions in primary healthcare: A systematic review of reviews. *Alcohol and Alcoholism*. 2014; 49:66–78. [PubMed: 24232177]
- R Core Development Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
- Roche AM, Freeman T. Brief interventions: Good in theory but weak in practice. *Drug and Alcohol Review*. 2004; 23(1):11–18. [PubMed: 14965883]
- Seale JP, Shellenberger S, Clark DC. Providing competency based family medicine residency training in substance abuse in the new millennium: A model curriculum. *BMC Medical Education*. 2010; 10(33):1–12. [PubMed: 20074350]
- Schwalbe CS, Oh HY, Zweben A. Sustaining motivational interviewing: A meta-analysis of training studies. *Addiction*. 2014 in press.
- Sullivan LE, Tetrault JM, Braithwaite RS, Turner BJ, Fiellin DA. A meta-analysis of the efficacy of nonphysician brief interventions for unhealthy alcohol use: Implications for the patient-centered medical home. *American Journal on Addictions*. 2011; 20:343–356. [PubMed: 21679266]

Fig. 1.

Group means and 95% confidence intervals for Motivational Interviewing Treatment Integrity scale of Routine Care Provider (RCP) and Non-Routine Care Provider (NRCP) Interventionist groups Motivational Interviewing Treatment Integrity coding system. Numbers on vertical axis identify individual interventionists. Vertical dashed and dotted lines represent Beginning Proficiency and Competency thresholds on the MITI, respectively.

Table 1

Descriptive Statistics for Routine Care Provider (RCP) and Non-Routine Care Provider (NRCP)
Interventionist Characteristics

	Interventionist Group		p-value
	RCP (n=9)	NRCP (n=6)	
Interventionist characteristics	<u><i>M (SD) / n (%)</i></u>		
Age ^a	45.9 (10.4)	32.3 (12.2)	.05
Women	7 (77.8)	3 (50.0)	--
White	5 (71.4)	3 (50.0)	--
Asian	2 (28.6)	0 (0.0)	--
Multiracial	0 (0.0)	3 (50.0)	--
Hispanic	2 (28.6)	1 (16.7)	--
Master's degree	9 (100.0)	3 (50.0)	--
Bachelor's degree	0 (0.0)	3 (50.0)	--
Years of clinical experience ^a	15.4 (7.5)	4.6 (7.2)	.02
Training dose and volume of work for the study ^a			
Workshop training hours	9.6 (1.7)	5.9 (2.4)	<.01
SP training sessions with feedback	3.2 (1.3)	2.3 (1.9)	.32
Months performing interventions for the study	18.3 (16.1)	14.9 (6.9)	.59
RP sessions performed with study participants	24.6 (21.8)	25.0 (11.4)	.96
RP sessions with feedback received from trainer	6.7 (4.5)	5.50 (2.4)	.57
Duration of intervention session (minutes)	22.06 (8.04)	26.03 (11.50)	<.05 ^b

Note. SP = Standardized patient training session; RP = Real patient brief intervention with study participant.

^a Between-group differences tested using independent samples t-tests.

^b Based on multilevel model controlling for nested data due to interventionist

Motivational Interviewing Treatment Integrity (MITI) Scores for Routine Care Provider (RCP) and Non-Routine Care Provider (NRCP) Interventionists

Table 2

MITI Scores	ICC ^a	Beginning Proficiency ^b	Competency ^b	Interventionist Group		z statistic	p-value
				RCP (n=9)	NRCP (n=6)		
				<i>M (SD)</i>			
% Open Questions	0.91	50%	70%	37 (21.0)	46 (20.0)	1.04	.43
% Complex Reflections	0.83	40%	50%	47 (16.0)	55 (15.0)	1.99	<.05
Reflection/Question Ratio	0.60	1.0	2.0	1.32 (0.79)	3.00 (3.90)	3.60	<.01
Spirit Global Score ^c	95%	3.5	4.0	4.41 (0.63)	4.68 (0.41)	1.95	.05
Empathy Global Score ^c	87%	3.5	4.0	4.55 (0.63)	4.89 (0.32)	3.55	<.01

Note. Calculations are based on 423 brief interventions (251 RCP, 172 NRCP). Between-group differences were tested using multilevel models including a random intercept for interventionist.

^a ICC = Intraclass Correlation Coefficients measuring inter rater agreement for three MITI coders, except for Spirit and Empathy (global 1–5 scales), for which is reported the percent of all pairwise coder comparisons that are within 1 scale point of each other.

^b Beginning proficiency & competency thresholds are preliminary expert opinions.

^c Global Likert ratings of Spirit and Empathy were strongly skewed and converted to binary variables (5 versus <5) for use in between-group analyses.