Routine outcome monitoring & learning organizations in substance abuse treatment
Oudejans, S.C.C.

Citation for published version (APA):
CHAPTER 3

FACILITATING & IMPEDING FACTORS FOR ROUTINE OUTCOME MONITORING (ROM) IN SUBSTANCE ABUSE TREATMENT

OUDEJANS, S.C.C.
NABITZ, U.W.
SCHIPPERS, G.M.
Abstract

To investigate facilitating and impeding factors and evaluate the support for routine outcome monitoring (ROM), we have asked key persons and professionals at four substance abuse treatment centers in the Netherlands to give their opinions in interviews and questionnaires on a project for ROM. The interview results show that the support of managers and treatment professionals is an important factor. The questionnaire results show that only a small percentage of professionals attend the feedback sessions, although most of them evaluate the feedback sessions as useful. Most professionals suggest continuing the project and over 80% are enthusiastic about the feedback sessions, thus warranting the conclusion that there is ample support for ROM. The interviews result in system-related recommendations that can easily be carried out, i.e. the desire that no additional questionnaires and differential methods are needed for data collection. To support quality improvement as a result of ROM, training is needed in the interpretation of the outcomes.
Introduction

The Dutch substance abuse treatment sector has made considerable efforts in the past ten years to introduce treatments that have proven to be effective and to establish transparency in the treatment results. The nation-wide quality project called Scoring Results was instrumental in this regard (Schippers et al., 2002). The introduction of cognitive behavioral therapy based on scientific evidence, called Lifestyle Training in the substance abuse treatment sector, is a convincing example of this. In the framework of the Benchmark Lifestyle Training in substance abuse treatment project, the results of Lifestyle Training are measured and fed back. For this purpose, routine outcome monitoring (ROM) has been set up at four substance abuse treatment centers. The aim is to continually improve the Lifestyle Training practice based on measurements of the results and thus achieve optimal results.

Interventions in health care and thus in substance abuse treatment as well are increasingly based on scientific evidence (W. R. Miller & Wilbourne, 2002; Mulder, 2004a, 2004b). Via scientific research, the efficacy of a treatment is tested under ideal conditions. Selecting the clients, for example without comorbidity, employing experienced clinicians, and devoting ample time and energy to protocol-training all help give these studies a high internal validity. The tried and true interventions are implemented in the day-to-day practice of substance abuse treatment, but the results from the scientific studies are not automatically achieved. In actual practice, groups of clients are usually more diverse and treatments are not always carried out according to the protocol (van Dijk, Schippers, & Visser, 2006). To determine the effectiveness of the treatment in daily practice, data is routinely collected using ROM. The aim is to increase effectiveness by improving the daily practice. There are two strategies for doing so. The first strategy involves feedback of outcomes of current treatments that is designed to make adjustments to the individual treatment process in progress. Individual treatment professionals receive feedback about individual clients, preferably related to norm data. This type of feedback gives treatment professionals information to support their work during the treatment. In the January 7 and February 8 issues of the Monthly Journal of Mental Health (Mgv), projects are described that apply this strategy (de Beurs & Zitman, 2007; Zwanepol & De Groot, 2008). The second strategy involves feedback of aggregated data on groups of clients afterwards. On the basis of these data, improvement plans can be formulated for treatment programs or treatment teams. In addition, data can be used for the purpose of accountability to internal and external stakeholders. Perhaps the term Routine Outcome Assessment (ROA) used by Wiersma and Sytema (Wiersma & Sytema, 2005)
would be appropriate here. This type of outcome assessment is a retrospective measurement and processes can no longer be adjusted at the individual patient level. To remain in step with the national and international literature, we use the term ROM, which can thus consist of two different strategies.

Not much is known about the effectiveness of either of these strategies. They cannot always be assumed to lead to improvements in the quality of care, but are the topic of a great deal of discussion and opinions are generally enthusiastic (Grol, Wensing, & Braspenninck, 2001; Schramade, 2005). This has in part been why the Health Care Inspectorate includes outcome data in the mandatory set of performance indicators (GGZ Nederland, 2006). The centers are thus accountable to the Health Care Inspectorate, medical insurance companies and society, and are expected to use the performance indicators for their internal quality policy.

The Amsterdam Institute for Addiction Research (AIAR) set up a ROM system for Lifestyle Training. The outcomes are measured at four substance abuse treatment centers and fed back at the aggregated level. Despite the enthusiastic responses at the centers and prior experience with the routine measurement of outcomes and performance data (Nabitz & Walburg, 2002), introducing ROM at the centers has been far more complicated than expected. There are a number of methodological (definition of criteria, inclusion of client groups), technical (extraction of data from the databases, presentation of data) and administrative problems (accuracy of the data, reaching clients after treatment) that have largely been solved. Setting up a good data collection system, feeding back the results and starting up an improvement cycle have all been more problematic than was anticipated. This led to additional tasks for treatment professionals and staff that were not always compensated, such as administering questionnaires and attending feedback sessions. The questionnaire response percentages and attendance at feedback sessions continued to be low. There was thus a vicious cycle of collecting low quality data that was unreliable, poorly distributed and was barely used for improving the quality of care. This pattern is not unique and this issue has also been touched upon in international literature (Ganju, 2006; Marsden et al., 2008; Teruya et al., 2006; Tiet et al., 2006).

This is why we decided to conduct further research into the factors that play a role in the introduction of ROM in the substance abuse treatment sector. The following two questions have been formulated for this purpose:
– What factors have facilitating or impeding effects when setting up and introducing a system for ROM?
– What is the support base amongst professionals for the ROM system and what suggestions do they have for improving it?
The research has been conducted amongst individuals who play a key role at the four ROM project centers or who carry out Lifestyle Training, either as treatment professionals or supervisors.

**Methods**

**Setting**

The research was conducted at four substance abuse treatment centers. At each center, approximately 2,500 new clients come in for treatment every year. Half of them are given treatment in the form of Lifestyle Training. This is consistent with the guidelines for patient allocation to appropriate levels of care (Merkx et al., 2007).

**Lifestyle Training**

Lifestyle Training is a type of outpatient treatment for substance dependence that was introduced around 2000. Based on cognitive behavioral therapy, it consists of registering the substance use, analyzing the advantages and disadvantages of substance use, learning to recognize and avoid high risk situations, and practicing alternative ways of regulating emotions and coping with craving. The treatment is focused on abstaining or regulating use and preventing relapses (W. R. Miller & Wilbourne, 2002). Four types of Lifestyle Training have been developed. Type 1 is for clients with mild, singular addiction problems. Type 2 is for clients with mild addiction problems involving various substances and/or co-morbid psychiatric problems. Types 3 and 4 are group variants of Types 1 and 2 (de Wildt, 2000a, 2000b; Merkx & van Broekhoven, 2003; van den Broek & Merkx, 2003).

**Routine outcome monitoring (ROM)**

In the ROM project *Benchmark Lifestyle Training in substance abuse treatment* (further referred to as the Benchmark Project), the Lifestyle Training outcomes at the four centers are continuously measured and retrospectively fed back at the aggregated level to treatment professionals, managers and Boards of Directors. At the start of the implementation, three measurements were employed, a baseline measurement during the intake, an exit measurement with a questionnaire on paper immediately after completion of the Lifestyle Training, and a telephonic follow-up measurement at 9 months after intake. During the course of the Benchmark Project, the structurally low response rate has led to the decision to remove the exit measurement. This was a questionnaire for patients, to be...
distributed to them by treatment professionals at the end of the Lifestyle Training. Baseline data are based on the regular intake interview. The telephonic follow-up measurement is done by a professional call center. Interviewers call clients nine months after intake and do an interview of about fifteen minutes.

The data about client features and substance use at the start comes from the electronic patient records, as does the data about treatment exposure. The follow-up data is from the call center database. The data is aggregated, analyzed and interpreted anonymously. The outcomes in terms of changes in alcohol and drug use, satisfaction and quality of life are linked to the client features and substance use at the start and to the nature (type of Lifestyle Training) and quantity of treatment exposure.

Every six months there is feedback to the treatment teams at meetings. There have been three of these meetings thus far and three semi-annual reports (Oudejans, Schippers, & Spits, 2007; Oudejans, Schippers, & Spits, 2006a, 2006b). After several multiparty sessions (benchmarks), there is now feedback within and to each participating center separately. The background and methodology of the Benchmark Project has been formulated in a manual (Oudejans & Schippers, 2006).

Data have been collected of approximately 4,000 clients thus far. It was possible to collect the personal and treatment exposure data of 95% of these clients. This has been the case with 60% to 95% of the baseline data on substance use and 53% of the clients have been reached for a follow-up interview.

Research design

The study consisted of two parts. In the first part of the study, key persons were interviewed about their experience with the Benchmark Project. In the second part of the study, all the professionals were given a questionnaire asking their opinions about the Benchmark Project.

Interviews with key persons

The key persons for the Benchmark Project were members of the Boards of Directors, the management, officials supervising the quality, team leaders and individuals responsible for the contents of the treatment.

The interview was based on a list of factors drawn up by the Netherlands Organization for Applied Scientific Research that play a role in implementing guidelines and treatment innovations in health care. The list has been revised into an interview guideline to access information on thirty-two factors (Fleuren, Wiefferink, & Paulussen, 2002; van Dijk et al., 2006). The respondents indicated
whether – on the grounds of their experience and perspective – each factor can facilitate or impede the implementation of a ROM project like the Benchmark Project. It was emphasized that the focus was on three aspects of the project: (1) the collection of data, e.g. paper questionnaires or electronic patient records, (2) the feedback of outcomes based on the data collected and (3) the improvement of treatment practice as a result of these outcomes.

All the questions in the questionnaire were discussed at the interviews and if the respondent indicated that a factor had a facilitating or impeding effect, he or she was asked to explain how. The interviewer was not involved in setting up the Benchmark Project, which assured independence during the interviews.

To be able to formulate a statement about the influence of the factors, we have performed calculations on twenty-eight factors that at least 60% of the respondents had an opinion about (i.e. being facilitating, impeding or having no influence). We concluded that the other four factors were not appropriately selected.

A factor of which more than 80% of the respondents had an opinion about in a certain direction (i.e. being facilitating or impeding) is said to exert an influence and is classified as an influential factor. A factor was viewed as either facilitating or impeding if it was influential and if more than 50% of respondents indicated that it has a facilitating or impeding effect. The respondents viewed some of the factors as having a facilitating as well as an impeding effect, thus indicating that the factor had two sides to it. The influence of these factors is more effective than the influence of a factor viewed as only having one side to it. The number of times a factor is said to have a facilitating effect added to the number of times it is said to have an impeding effect is what we call the impact.

Questionnaire for professionals

In the second part of the study, a questionnaire was sent to all the treatment professionals administering the Lifestyle Training and their supervisors at the centers participating in the Benchmark Project. By supervisors, we mean the team leaders, management and individuals responsible for the contents of the treatment. Three feedback and benchmark sessions had taken place when the questionnaires were sent to the treatment professionals.

The questionnaire consisted of six questions about the Benchmark Project and was developed by the Benchmark Project team. The teams’ staff members evaluated its face validity. The questions pertained to whether or not respondents had attended the feedback sessions and the evaluation of these sessions including a grade from 1 to 10. The respondents could also make suggestions for improvements. For each question, the percentage of respondents that gave a certain answer was calculated. Differences between centers and treatment professionals and supervisors were also examined.
Results

Respondents: Key persons and professionals

Twenty-one key persons have been approached. Nineteen of them (90%) have been interviewed. The analysis of the various centers separately showed that all of them were represented by the same number of key persons, and in the same way. Table 1a shows the positions of the respondents at each center.

A total of 170 questionnaires were sent to professionals, and 101 of them (59%) were filled in and sent back. Eighty-nine came back in the first round and another twelve after a reminder was sent. Table 1b shows the features of the respondents. The response percentages at the centers vary from 37% to 92%. Three quarters of the responses are from females, most respondents have a B.A. or M.A., and 87% are treatment professionals and 13% supervisors. The respondents at the various centers do not differ significantly on these features, although a higher percentage of respondents at one of the centers have an M.A. The respondents who filled in and returned the questionnaire in the first round do not differ from the twelve respondents who did so after receiving a reminder.

Table 1a:
Key persons characteristics (n = 19)

<table>
<thead>
<tr>
<th>Position</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Board of Directors</td>
<td>21</td>
</tr>
<tr>
<td>Official in charge of Benchmark project</td>
<td>32</td>
</tr>
<tr>
<td>Team leader</td>
<td>21</td>
</tr>
<tr>
<td>Responsible for treatment contents</td>
<td>21</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1b:
Professionals’ characteristics (n = 101)

<table>
<thead>
<tr>
<th>Position</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifestyle trainer</td>
<td>87</td>
</tr>
<tr>
<td>Supervisor</td>
<td>13</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>25</td>
</tr>
<tr>
<td>Female</td>
<td>75</td>
</tr>
<tr>
<td>Age (average sd)</td>
<td>41.2 (11.4)</td>
</tr>
<tr>
<td>Educational level*</td>
<td></td>
</tr>
<tr>
<td>4, 5 or 6 year secondary school or less</td>
<td>2</td>
</tr>
<tr>
<td>College (B.A.)</td>
<td>60</td>
</tr>
<tr>
<td>University (M.A.)</td>
<td>31</td>
</tr>
<tr>
<td>Other (Post-college/university)</td>
<td>7</td>
</tr>
</tbody>
</table>

* significant difference between the centers ($\chi^2_{9} = 27.7, p = 0.00$)
Results of interviews with key persons

Fourteen factors are viewed by more than 80% of respondents as being influential (Table 2). The percentage of facilitating factors has been calculated by adding the percentage of respondents who indicate that a factor has a facilitating effect to the percentage of respondents who indicate that the factor can have a facilitating as well as an impeding effect. The same procedure is adhered to for the calculation of the percentage of impeding factors. The percentages of facilitating and impeding factors can consequently add up to more than 100%.

Table 2 shows that ten factors were viewed by more than half the respondents as having a facilitating effect. The factors viewed by the largest number of respondents as having a facilitating effect are Support for the project of supervisor at the center and Adequate material facilities (questionnaires and so forth). A total of 92% of respondents viewed both these factors as having a facilitating effect.

Of the 14 factors, two were viewed by more than half of the respondents as having an impeding effect. The factor viewed by the largest number of respondents as being impeding is Insufficient time available for filling in the questionnaires. A total of 85% cited this as having an impeding effect.

The factors the respondents viewed as having a facilitating as well an impeding effect apparently have two sides to them and have a large impact. The
five factors with the largest impact are *The feeling that the project belongs to the people involved*, *The explicit opinion of one or more of the people involved (opinion leader)*, *Support for the project from colleagues at the center*, *Adequate material facilities* and *The commitment of treatment specialists in setting up the project*.

The fourteen factors include five that pertain directly to the sub-topics data collection, feedback and improvement. As regards data collection, *Adequate material facilities*, *Adequate administrative support* and *Sufficient time available for filling in the questionnaires* are influential factors. The third one is markedly impeding and the other two have a facilitating effect. As regards feedback, *Attractive feedback sessions* is viewed as influential, and as regards improvement, *The treatment professional is able to interpret the outcomes* was viewed as influential and facilitating.

Results of the questionnaire for professionals

A total of 35% of the 101 respondents have attended the Benchmark Project feedback sessions. They answered three questions evaluating the feedback sessions. The large majority (87%) felt the sessions are relevant to the work and 94% felt the center should continue with this kind of feedback. The respondents graded the feedback sessions with an average of 7.1 ($n = 34$, $sd = 0.74$) and almost 80% gave a grade of 7 or higher. No unsatisfactory of 5 or lower grades were given (Figure 1). There are no differences in opinion about the Benchmark Project between the treatment professionals and supervisors.

![Figure 1: Professionals' opinions on Benchmark Project](image)

All 101 respondents were asked whether comparing the outcomes with those at other centers was useful in the course of their work. The large majority of respondents (81%) indicated that comparing the outcomes is indeed useful
in the course of their work. We checked to see if there was any correlation between their answers to this question and whether or not they attended the feedback sessions. No such correlation was observed.

Differences were observed though between the respondents at the various centers as regards the feedback session attendance, the feedback relevance evaluation and the views on whether the feedback sessions should be continued. Time and again, respondents at one of the centers appear to be the most critical. Almost 73% of respondents at one of the centers appeared to have attended one or more feedback sessions, though this was only true for 17% of respondents at another center ($\chi^2(3) = 13.8$, $p < 0.05$). This is also the center where the respondents were most critical about the relevance of the feedback, since only 40% indicated that they found the feedback relevant. This is true for 100% and 88% at the other centers. The respondents at this center were also the most critical about continuing the feedback sessions, since only 60% indicated that they are in favor of continuing. This is true of 100% at the other centers ($\chi^2(3) = 11.5$, $p < 0.05$).

A total of 31 of the 101 respondents wrote down suggestions for the outcome collection and feedback. The largest number (12) suggested some other or supplementary feedback, i.e. sending a digital form of the outcomes to all the professionals personally, putting the results on the Internet or making a concise form of the outcomes available on one or two printed pages.

Ten of the respondents gave comments or suggestions about how the implications of the feedback are dealt with at the centers. The outcomes remain invisible at the centers and it is unclear to the professionals how improvement actions can be launched on the basis of the results. They also commented that the supervisors do not devote much attention at all to the results. Three of the respondents have suggestions about the planning of the feedback sessions. They are either announced too late or the announcements are poorly distributed at the centers. Making attendance mandatory for managers was also mentioned. It was also noted that a central feedback session requires a lot of traveling time and valuable production time so that it would be better to plan the feedback sessions at the various locations.

**Conclusions and discussion**

Using interviews and questionnaires, we have focused on factors that play a role in the introduction of ROM and its support base. Fourteen factors emerged that are found to be important. It is facilitating if there is support and commitment on the part of staff and supervisors, if the outcomes are comprehensible and if they are attractively presented. Commitment is also mentioned as an
impeding factor though, i.e. if the respondents do not have the feeling the project is *theirs*. So in launching ROM it is important for as many people as possible to be aware of its importance so there is a sturdy support base. Another concrete impediment to ROM projects is a lack of time for filling in questionnaires. Together with making sure there are enough questionnaires, which has a facilitating effect, this is a practical matter. Which is not so surprising, but it does make it very clear that time and money need to be set aside for practical matters, otherwise it will be hard to put ROM into effect.

If we look at the number of times various factors are mentioned as having either a facilitating or an impeding effect, we see that *commitment* and *support* are the leading ones. This is why it is important that the questionnaire shows that this support is there for the Benchmark Project on the part of treatment professionals as well as the supervisors.

Most of the professions see the relevance of feedback. In addition, the feedback sessions are appreciated and most of the professionals feel that data feedback in this form should be continued. The attendance at the feedback sessions is limited however. So for the time being, the distribution of the outcome data is limited as well.

The lack of time to devote attention to filling in questionnaires and the lack of administrative support were facts that were mentioned as having an impeding effect. As regards the Benchmark Project, this manifested itself in a low response (sometimes only 25%) at the exit meeting, in our case the second measurement. A low response and a big burden on staff is not an unknown phenomenon at ROM projects where the treatment professional plays a large role in administering the questionnaires (Harrison & Asche, 2001; Teruya et al., 2006; Zwanepol & De Groot, 2008). This result has led us to remove this measurement. By extracting data for the baseline measurement from the intake interview and having the follow-up measurement done by a professional call center, we have seen to it that no additional questionnaires are required during the treatment process and that extra activities on the part of the treatment professionals are no longer required for the ROM data collection. This way of working guarantees a sufficient response, no extra administrative burden for the professionals and an independent assessment of the treatment outcomes.

There is a striking difference between the opinions of respondents at the various centers. It is mainly the respondents at one center who were less enthusiastic about ROM. The criterion seems to be how often they attended the feedback sessions. The attendance figures and interest level were low at this center and the opinion about ROM was less enthusiastic. Perhaps these professionals first needed to be motivated to accept ROM by witnessing the advantages at other centers.
To increase the attendance at feedback sessions, we have since switched from one central semi-annual feedback session to semi-annual visits to each of the centers. In addition, it is important to also distribute the data via the Internet, posters or e-mail.

The dearth of improvement actions resulting in an observed improvement of the substance abuse treatment – the aim of ROM – as is confirmed in the questionnaire study by the suggestions made by the professionals, has been an important reason for this study. This is also the picture that emerges from the literature. Up to now, the expectation that measuring what is done will lead to increased information and a better quality of substance abuse treatment, has not been fulfilled (Ganju, 2006). Professionals indicate that the outcomes are not dealt with effectively at the centers. The supervisors do not devote much attention to them and it is not clear to them how they can launch improvement actions on the basis of the results. So there is a lack of support on the part of the supervisors and of general commitment to the project. The essential factors that could have a facilitating effect are not in evidence. This is consistent with findings from the United States: in the California Treatment Outcome Project (CalTOP), lack of staff buy-in was - together with the lack of sufficient funds and resources - one of the dominant factors as mentioned by key persons and stakeholders (Teruya et al., 2006). Regular overviews of the outcomes at team level or giving teams an extra say in the form and contents of feedback could be effective in this connection. Addressing and interpreting treatment results in training and other courses could lead to more improvement actions and thus improve the quality of the treatments in such a way as to reach the efficacy of the RCTs the treatments are derived from.

This study has its limitations, such as the response to the questionnaires for professionals. It is true that almost 60% of them have filled in the questionnaire and sent it back, but there is the risk of distortion because perhaps only the professionals who are familiar with ROM make the effort to fill in the questionnaire. The fact that no differences in the results are observed between the professionals who sent back the questionnaire immediately and those who did so after the reminder, is an indication that the risk of distortion is limited. In addition, we have only examined one form of ROM, i.e. retrospective feedback of aggregated information. This does not necessarily mean these results also hold true for the ROM strategy with feedback on current treatments. It is not uncommon for the data used for this purpose to also be aggregated for the evaluation of treatment programs and actions to improve their quality (de Beurs & Zitman, 2007; Zwanepol & De Groot, 2008), so that at any rate our results are valid for this application.

The introduction of a nation-wide system for performance indicators in
the mental health system in general and substance abuse treatment in particular (GGZ Nederland, 2006) makes the improvement of substance abuse treatment on the basis of the results all the more relevant. Centers make sizable investments in collecting the data for these indicators and will be held accountable for them in the future. So it is important that teams are backed by quality departments that support them in setting up ROM and translating the results into improvements. There is enough of a support base for this.

Acknowledgements

We would like to express our gratitude to the management, the key persons and the treatment professionals at the various centers for taking part in the study. The Jellinek (now the Jellinek division of Arkin), Brijder Substance Abuse Treatment Center (now the Brijder Substance Abuse Treatment Division at ParnassiaBavo Group), and Novadic-Kentron have made it possible for us to conduct the study. We also thank Judith Noijen for doing the interviews.