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

Citation for published version (APA):
Extent and costs of substance abuse

There is a growing emphasis on quality and results of health care organizations, and this is also true for the substance abuse treatment sector. The abuse of alcohol and other drugs is associated with serious public health problems, public safety problems and economic damage. In the United States, 5% of people aged over 12 meet the criteria for substance dependence (Dennis & Scott, 2007). In Europe, 12 month prevalence ranges from 1.1% in women to 6.1% in men (Rehm, Room, van den Brink, & Jacobi, 2005; Rehm, Room, van den Brink, & Kraus, 2005), and according to the most recent estimations for the Netherlands, 8.2% of the adult population meets the criteria for alcohol abuse or dependence (3.7% meet the criteria for dependence) and about 1.3% for drug abuse or dependence (Verdurmen, Monshouwer, van Dorselaar, & de Graaf, 2003; Vollebergh et al., 2003), representing around 800,000 adults with alcohol abuse or dependency and 130,000 adults with drug abuse or dependency.

Costs of substance abuse are high. The estimated tangible costs of harmful alcohol use to the European Union (EU) for 2003 are estimated at 125 billion euros. Spending on alcohol-related problems like health care, the criminal justice system and traffic accidents accounted for 61 billion euros. The costs of potential production not realized due to absenteeism, unemployment, and premature mortality were estimated at 59 billion euros. The remaining 5 billion euros were spent on treatment and prevention (Cnossen, 2006). Tangible costs of harmful alcohol use in the Netherlands were estimated at a minimum of 0.7% of the GDP. This is low in comparison with other EU member states such as Belgium (1.7%), but higher than Italy and Portugal (both 0.4%). Taking into account a GDP of 600 billion euros, costs for the Netherlands are at least 4.2 billion euros. In all EU member states, except Finland, excise duty collections fall short of the estimated tangible costs (Cnossen, 2006).

These costs, combined with serious public health and public safety problems, and the fact that most treatments for substance abuse are provided through public funding, mean there is a legitimate interest of stakeholders like governments, health insurers, and consumers of care to assure that substance abuse treatment centers are effective and efficient, i.e. deliver high quality care for a reasonable price. In addition, stakeholders put pressure on the sector to demonstrate programmatic transparency and accountability regarding treatment outcomes.

1 Taking into account a population of 10 million adults (source: Statline, CBS (http://statline.cbs.nl/StatWeb/publication/?dm=SLNL&PA=37296&ED=1=0-5&ED2=0,10,20,30,40,50,(l-1)-l&VW=rt)

Demonstrating, maintaining, and improving quality of health care requires performance measurement, measurement of treatment outcomes, and learning from feedback.

Quality of health care

The Institute of Medicine (IOM) defines quality of health care as: “The degree to which health services for individuals and populations increase the likelihood of desired outcomes and are consistent with current professional knowledge” (Institute Of Medicine, 2001). In 1966, Donabedian proposed a three-element model for quality measurement in health care, consisting of structure (characteristics of the health care setting), process (what is done), and outcome (the status of a patient after a series of interventions) (Donabedian, 2005; Loeb, 2004). With his model, he established the current field of performance management and outcomes research. However, measuring quality of care originates further back than Donabedian, and two historical names in that perspective are Florence Nightingale and Ernest Codman. During the Crimean war in the early 19th century, Florence Nightingale collected mortality data and infection rates for hospitals in England, resulting in a decrease of soldiers dying from bad living conditions in hospitals. It is said she was a pioneer in the visual presentation of information and statistical graphics at the military field hospital she managed (for instance to illustrate seasonal fluctuations of patient mortality). Decades later, in the early part of the 20th century, Ernest Codman developed what is known today as outcomes management in health care, by proposing a measurement system for results of care of surgeons (Loeb, 2004). According to Codman’s “End Results Idea”, every hospital should “follow every patient it treats, long enough to determine whether or not the treatment has been successful, and then to inquire ‘if not, why not?’ with a view of preventing similar failures in the future” (Kaska & Weinstein, 1998). In Codman’s End Results Hospital each patient was provided with an “end result card” on which the surgeon filled out the details of the case before and after surgery, as well as the symptoms, the diagnosis, the treatment plan, complications that occurred, and the diagnosis at discharge. A year later the “end result card” was brought up, and based on the patient’s current condition the treatment was evaluated. Results were put on the end result card each year afterward, until a definitive determination of the result of treatment could be made. Codman also believed that the information should be public, which enabled patients and providers to compare results of various treatments among surgeons and different hospitals (Kaska & Weinstein, 1998).

Over the years, many definitions of quality of health care have been suggested. In many health care sectors, patients as well as governments and health insurers have become “consumers”, and doctors and treatment counselors have become “providers”, implying an important change in the roles of the players in the field (McLellan, Chalk, & Bartlett, 2007). McLellan proposes a practical and usable definition of quality care as: “evidence-based treatments that are provided by licensed or credentialed practitioners who have demonstrated core competence in their practice areas and whose activities are monitored regularly by program- and system-level measurement of quality indicators” (McLellan et al., 2007, p.334). In this definition, performance or quality indicators establish the extent to which the care being delivered conforms to evidence-based practices. In addition, the definition can be extended to adherence to best practices, if no evidence-based practices are present. However, this definition does not state anything about the actual contents or aspects of care that are measured with the performance indicators. In that respect, the six core needs for health care as proposed by the IOM are very useful. In The Quality Chasm, a report in which aims for a new health system for the 21st century are laid out, the IOM presents six core needs for health care: it should be safe, effective, patient-centered, timely, efficient and equitable (Institute Of Medicine, 2001). It is on these core needs that organizations can provide performance indicators.

There is, however, debate about whether treatment outcome should be regarded as a performance indicator. Perrin argues that outcome is only meaningful as a measure of performance if there is a causal relationship between the structure and process of care on the one hand and the outcome on the other hand (Perrin, 2002). The structure and the process are the performance itself and the outcome should then be the logical result of that performance. Therefore, McLellan puts measures of effectiveness and outcomes aside from performance indicators (McLellan et al., 2007). Sperry does the same in separating process evaluation from outcomes evaluation (Sperry, Brill, Howard, & Grissom, 1996). We prefer to use the model of Donabedian, where there are indicators on the way the treatment is delivered (structure and process indicators) as well as measurements of the results of the treatment (outcome). In this, the IOM core need of effectiveness is the result or the outcome, and the other five core needs are to be considered as process indicators. Measuring, maintaining and improving quality of health care through performance indicators is often described in systems for quality management.
Quality management in health care

During the last decades, most health care providers implemented some type of quality management system by which they can identify, measure, control, and improve the various core business processes that should lead to improved business performance. The most common term in this field is Total Quality Management (TQM). TQM is a management strategy aimed at embedding awareness of quality in all organizational processes. TQM was first promoted in Japan with the Deming prize, and was adopted and adapted in the USA as the Malcolm Baldrige National Quality Award and in Europe as the European Foundation for Quality Management award (EFQM). Both the Baldrige Award and the EFQM Model comprise of a set of standards against which the features of the organization are examined. The EFQM Model (Figure 1) consists of five areas on organizational structure: Leadership, People, Policy and Strategy, Partnerships and Resources, and Processes. These areas provide structure to describe the way an organization functions, they also function as a tool to guide organizations towards better performance. The model contains four further areas on results of the organization: People Results, Customer Results, Society Results and Key performance Results. For every result area the organization determines what measure is relevant, how the measures will be employed and what is done with the results of these measurements. The feedback loop of Innovation and Learning implies that the organization has to learn from previous results to stay on track, to improve or to innovate. Empirical studies on the effectiveness of the EFQM approach are scarce (Nabitz, 2006), but single case studies suggest the approach can be successful in broadening quality awareness in teams and in the implementation of evidence-based treatments in organizations (Nabitz, Schaefer, & Walburg, 2006; Nabitz, Schramade, & Schippers, 2006).

Figure 1: The EFQM Model
Quality improvement by learning from feedback and benchmarking

In management science it is assumed that quality improvement in organizations profits from performance measurement and feedback of results. This is also recognized in health care management, by adopting quality systems such as TQM or the EFQM Model where innovation and learning from previous results are central concepts (Colton, 2000; Loeb, 2004; Nabitz, 2006; Thier & Gelijns, 1998). Generally, feedback of results is employed in two strategies: The first strategy is to feed back results during the treatment process, aiming at adjusting the concurrent process if necessary. This type of feedback is to support individual treatment professionals in their decision-making. The second strategy is the feedback of aggregated results at the end of the process. The aim is to formulate improvement projects for treatment programs or other processes in the organization, or to provide stakeholders with information on accountability. However, the effects of both strategies of feedback of performance on professional practice and health care outcomes are not clear-cut. The Cochrane Institute issued reviews on this topic, and they conclude that audit and feedback can sometimes be effective in improving the practice of health care professionals. The effects seem to be small to moderate, but potentially worthwhile (Jamtvedt, Kristoffersen, O’Brien, & Oxman, 2006; Jamtvedt, Young, Kristoffersen, O’Brien, & Oxman, 2006; Thomson O’Brien et al., 2000, 2003).

A specific type of feedback provision is benchmarking. Benchmarking is a process used in management in which organizations compare various aspects of their processes in relation to that of the best practice, usually within their own sector. Established reference points are used to interpret data. These reference points can be best practices or external results, like results from randomized clinical trials (Mulder & de Loor, 2005; Walburg, 2001).

Routine outcome monitoring

A specific approach to improve clinical practice through feedback and benchmarking is routine outcome monitoring (ROM), where the focus is on measuring and reporting on health outcomes. ROM is a method to assess outcome of treatments by measuring the nature and severity of patients’ symptoms periodically, basically according to Codman’s End Result Card. This approach

---

4 Jamtvedt defines audit and feedback as “any summary of clinical performance of health care over a specific time period”, therefore audit and feedback will be treated as one entity, although both terms are also defined separately.
is strongly advocated by Ellwood. He proposed a “common patient-understood language of health outcomes, a national data-base containing information on clinical, financial, and health outcomes, that estimates as best as we can the relation between medical interventions and health outcomes, as well as the relation between health outcomes and money; and an opportunity for each decision-maker to have access to the analyses that are relevant to the choices they must make” (de Beurs & Zitman, 2007; Ellwood, 1988). Ellwood also noted that outcome management would be a “clinical trial machine”, i.e. a routine part of medical care that would never stop, and standards and measures would be constantly modified based on feedback. The idea was welcomed, and more publications emerged stating the importance of outcome measurement, the creation of routinely accessible “real-world” data sets, and the elimination of undesirable variations in clinical practice (Holloway, 2002). Outcomes that should be measured in ROM are the functioning and well-being of the patient, together with disease-specific outcomes. Sperry distinguishes clinical outcomes and functional outcomes (Sperry et al., 1996). Clinical outcomes are those that describe the physical and psychiatric signs and symptoms of a disease or a disorder, whereas functional outcomes are those that describe levels of patient functioning in such areas as work, family, health and grooming, intimate relations, self-management, and social relations. Others propose to extend the set of measures to patient satisfaction, costs of treatment, or numbers of readmissions as part of the ROM system (Thier & Gelijns, 1998; Walburg, 2001).

Strategies in ROM are similar to the two strategies in the feedback of results. The first strategy in ROM is to feed back outcome during an individual treatment process, with the aim of adjusting individual treatments if necessary, as is illustrated with the grey lines in Figure 2. In this strategy, treatment professionals receive information about their individual patients, preferable related to existing norms. This type of feedback is to directly support individual professionals in decision-making. For instance, several studies in mental health care showed that outcome of poorly responding patients in an early stage of treatment can improve with feedback on outcome (Lambert, Harmon, Slade, Whipple, & Hawkins, 2005; Sapyta, Riemer, & Bickman, 2005). The second strategy is the feedback of aggregated results over groups of patients afterwards, as is illustrated with the black lines in Figure 2. The aim of this strategy is to formulate improvement projects, not for individual patients, but for treatment modalities or programs. In addition, aggregated figures can serve as accountability figures to internal or external stakeholders (Lambert et al., 2005; McKenzie & Marks, 2003; McLellan et al., 2007; Nabitz & Walburg, 2002).
Sperry provided definitions for outcome monitoring and outcome management (Sperry et al., 1996). He defined outcome monitoring as: “The use of periodic measurement or assessment of treatment outcomes over time compared against expected outcomes in order to alter treatment, compare treatment interventions, or make attributions about what produced change. Typically, feedback of data to the provider serves as the basis for modifying the plan and the course of treatment”, thereby classifying the term “monitoring” as the appropriate term for the concurrent, individual feedback strategy. Outcome management is defined as: “The use of monitoring data in a way that allows learning from experience. Usually this results in reshaping or improving the administrative and clinical processes of services provided. Patient profiling and provider profiling are additional aspects of outcome[s] management systems”, using the word “management” for the aggregated feedback strategy.

Routine outcome monitoring (ROMon) therefore refers to the concurrent,
individual feedback strategy and routine outcome management (ROMan) to the aggregated feedback strategy. Although the terms ROMon and ROMan are attractive from an educational point of view, we will use the term routine outcome monitoring (ROM) for both strategies with reference to one of the two strategies – the concurrent, individual (monitoring) strategy or the aggregated (management) strategy. This choice is partly conceptual and partly practical. As Sperry’s definitions imply, data collected for outcome monitoring are frequently also used for outcome management and both activities are often intertwined, justifying the use of a term that encloses both strategies and activities. In our experience, “monitoring” covers both strategies best. In regard to the practical choice, the distinction between “monitoring” and “management” vanishes with the convenient use of the abbreviation ROM. Therefore; we chose to use one meaning for the letter “M” in ROM term, providing an addition to the term to clarify the used strategy when needed.

**ROM in mental health care and substance abuse treatment**

As early as 1990, the Institute of Medicine described the principles of outcome monitoring in the substance abuse treatment sector (Institute Of Medicine, 1990). Sperry contributed a volume on ROM in mental health care and substance abuse treatment, in which several ROM projects were presented as illustrative cases (Sperry et al., 1996). Gradually, ROM projects and publications started to appear in mental health care (Holloway, 2002; McKenzie & Marks, 2003; Sperry et al., 1996) and in the area of substance abuse treatment (Evans & Hser, 2004; Harrison & Asche, 2001a; Moos, Finney, Federman, & Suchinsky, 2000; Tiet, Byrnes, Barnett, & Finney, 2006). These studies were mainly focused on evaluating the service delivery system and enhancing treatment outcomes. In all cases, the assumption is that programs or organizations will profit from ROM through learning from the feedback of outcome data.

**Learning organizations**

A concept related to feedback and benchmarking of performance and outcomes is the “learning organization”. In Scotland and England, the National Health Services have been encouraged to become learning organizations to meet the demands of reform and change in health care (Kelly et al., 2007; Rushmer et al., 2007) and it is stated by Ellis that a learning organization is required for benchmarking to be effective (Ellis, 2006). Learning from experience was also mentioned in Sperry’s definition of outcome management. According to the organization model of Senge (Senge, 1992), learning capacity is associated
with organizational effectiveness, and the learning principles of shared vision and team learning are predictors for organizational effectiveness (Jeong, Lee, Kim, Lee, & Kim, 2007; Kelly et al., 2007). Since the theory of Senge incorporates the role of feedback, it is very applicable in the field of quality improvement through ROM.

In his book *The Fifth Discipline*, Senge defined a learning organization as one that possesses five core-learning disciplines: Personal Mastery, Mental Models, Shared Vision, Team Learning, and Systems Thinking (Senge, 1992). Systems Thinking is stressed as the core discipline that integrates the others into a unity of theory and practice. Systems Thinking is meant as “the fifth discipline”. Another essential discipline is Team Learning, because this is a requirement for organizational learning: only when teams have the ability to learn, is organizational learning attainable. The combination of these five disciplines is a permanent course of practice for individuals, teams, and organizations. The five disciplines of a learning organization are (Senge, 1992):

1. **Personal Mastery** – the process of expanding personal capacity and continually improving one’s level of expertise in order to achieve desired goals.

2. **Mental Models** – internal pictures, assumptions, and generalizations of the world that influence perceptions, reactions and decisions.

3. **Shared Vision** – the capacity to develop commitment to and in an organization or a group, by establishing the principles and guiding practices to develop shared pictures of the future desired by members.

4. **Team Learning** – the capacity of groups to acquire collective thinking skills through dialog, in order to develop intelligence and results that wouldn’t be reached individually.

5. **Systems Thinking** – the ability to use knowledge and instruments in order to perceive the greater patterns in systems and organizations and to change them effectively.

The theory of Senge implies that learning organizations are expected to book better results than organizations with less learning capacity. Senge emphasizes the role of feedback in the learning organization. ROM therefore plays an important role in the learning organization and quality improvement through feedback and benchmarking of performance and outcomes.

**Aim of the thesis**

Policymakers, managers, and clinicians in the sector of mental health care and substance abuse treatment endorse the merits of ROM. However, from multiple implemented projects and available reports and studies on ROM in scientific literature, some unsolved issues have emerged. It is the aim of this
thesis to elaborate on these issues and to suggest possible solutions. Firstly, there is the issue of the feasibility of data collection and the validity of the collected data for ROM. Data collection can be expensive and time consuming (McKenzie & Marks, 2003; Tiet et al., 2006). As a result, response rates can be very low and results may be of questionable validity for the target population. Up to 20 interviewer hours per patient are reported, inclusion of patients can be very difficult and response rates can be as low as 48% (Gerstein & Johnson, 2000; Harrison & Asche, 2001a; Rosenheck & Seibyl, 2005). This is especially problematic in the case of aggregated ROM data, but also for individual concurrent feedback, since low response rates imply that only few patients will profit from this quality improvement. Secondly, while researchers, policymakers, and top managers generally support ROM projects; this support is often much lower under treatment professionals and lower-management staff (Harrison & Asche, 2001a; Teruya, Hardy, Hser, & Evans, 2006), resulting in low credibility of the ROM system. Thirdly, databases developed for ROM can be a rich source for add-on research on effectiveness of substance abuse treatment. But, in many cases, the methodology of studies in the field of quality lack precision in operationalization and standardized instruments (Nabitz, 2006). The question is whether it is possible to retrieve information from these databases to perform research on effectiveness on the database – we will provide an example of this. The fourth aim of this thesis is to elaborate on feedback and the learning organization. It is believed that the availability of information on treatment outcome, on the individual real-time level as well as on the aggregated level enhances the quality of treatment, although results of studies are not clear-cut and effects vary widely (Jamtvedt, Young et al., 2006; Thomson O’Brien et al., 2000, 2003). Therefore, an appropriate question is whether delivering feedback has an effect on the prerequisite for quality improvement, i.e. on the learning capacity of individuals, teams and organizations. In other words, does ROM have an effect in terms of Senge’s concept of learning organizations? These are questions on the effectiveness of feedback and benchmarking.

The questions in this thesis are concentrated around implementing ROM in substance abuse treatment and an attempt to answer the question: Is it possible to implement a kind of ROM that is feasible, generates valid data, is supported by key persons in services, can be used to generate conclusions about effectiveness and, finally, will contribute to the learning capacity of employees?

In answering the research questions, this thesis is limited to the ROM strategy of collecting data for feedback of aggregated results over groups of patients afterwards. This means that results and conclusions of this thesis are of restricted value for ROM projects aimed at the strategy of individual feedback.
on the concurrent treatment process. However, since it is not uncommon that these data are also used to report on aggregated levels (de Beurs & Zitman, 2007; Zwanepol & De Groot, 2008), this thesis is also of value for that type of application. The current research is performed within the setting of the substance abuse treatment sector in the Netherlands. This sector is part of the mental health care sector and most organizations are members of the branch organization for Dutch Institutions for Mental Health Care and Addiction Services (GGZ Nederland). Although research is carried out in the substance abuse treatment sector, we expect the findings to be relevant for the entire mental health sector.

Setting of the thesis

The Netherlands has an extensive network of medical and psychosocial treatment facilities for people with substance use disorders or pathological gambling. The centers are part of the mental health services. In 2006, a total of 60,976 patients were registered in Dutch outpatient and inpatient facilities. Almost 50% of them had a primary diagnosis of alcohol dependency, 21% had a diagnosis of opiate dependency, and 16% and 11% were dependent on cocaine and cannabis, respectively (van Laar, Cruts, Verdurmen, van Ooyen-Houben, & Meijer, 2008). Funding of the alcohol and drug treatment services is public. The costs for substance abuse treatment in 2000 amounted to around 125 million euro’s (ca. US$180 million), 51% of which was spent on outpatient services and 49% on inpatient services.

The definition of quality of care of the IOM implies that desired health outcomes should be achieved and current knowledge should be applied: this was the exact reason for a major reorganization that took place in Dutch substance abuse treatment services during the last decade. It was partly in reaction to public criticism that substance abuse treatment was not effective and not contributing to solving social problems and public nuisance. As a consequence, the sector had a serious image problem and treatment services suffered some serious drawbacks (Nabitz, Oudejans, Brink van den, & Vis, 2006; Nabitz, Vis, & Brink van den, 2001; Schippers, Schramade, & Walburg, 2002; Schippers, van Es, Mulder, & Dijk van, 2005). Therefore, the sector started a nationwide government supported reform called “To Score Results”. The objective was to redesign existing treatment services on the basis of scientific evidence regarding efficacy and effectiveness and to improve treatment practices on the basis of feedback on clinical and societal outcomes (Schippers et al., 2002). The existing treatment services were reorganized into services with a central intake location where protocolized treatment-allocation took place to evidence-based
interventions. However, the implementation of a system for the improvement of treatment practices based on feedback of outcomes, lagged behind. To boost this part of the project, the ROM project Benchmark Lifestyle Training in substance abuse treatment was started. In this project, outcomes of an outpatient treatment for substance dependence at four treatment centers are measured routinely and retrospectively fed back and benchmarked at the aggregated level to treatment professionals, managers and Boards of Directors. Next, we formulated research questions concerning implementation, feasibility and effectiveness of the ROM project. This thesis is a result of this research.

Overview of the thesis

Chapter 2 deals with the topic of feasibility and validity of data collected in ROM. It primarily describes the feasibility of a telephonic follow-up interview for ROM. Post-treatment follow-up interviews, designated to monitor and evaluate outcomes after treatment has ended are an important part of ROM. However, follow-up interviews are costly and time consuming. Investments of up to 20 interviewer hours per patient are reported in the literature, and response rates are generally (much) lower than 70% (Gerstein & Johnson, 2000). Therefore, we implemented a telephonic follow-up interview to conduct these interviews and studied the feasibility in terms of response rate, time investment and costs. Next, we investigated whether the collected data showed validity in terms of tracking difficulty and representativeness. We also studied the relationship between response-related baseline characteristics and treatment outcome. This was in order to assess whether the data suffered from selection bias, which would limit the value of generalizing outcomes from the follow-up sample to the total treated population.

Chapter 3 describes facilitating and impeding factors and the observed support for routine outcome monitoring (ROM) in substance abuse treatment centers in the Netherlands. During implementation of ROM we encountered problems in data collection and in the participation in feedback sessions. In order to identify facilitating and impeding factors and support for the system, key persons and professionals from four substance abuse treatment centers in the Netherlands were interviewed and filled out questionnaires to find out their opinions on the ROM project.

Chapter 4 is a naturalistic outcome study. It is an example of how databases for ROM can be utilized to conduct effectiveness studies. Such studies are a necessary addition to randomized controlled trials (RCTs) with their highly-selective patient samples and their overly-structured treatments. In RCTs, efficacy is tested, i.e. what are the results of a treatment under ideal conditions?
As a consequence, most if not all RCTs have low external validity and do not offer a realistic estimate of effectiveness, i.e. how well a treatment works in the real world of day-to-day clinical practice. Therefore, naturalistic outcome studies are performed in order to evaluate the merits of interventions in a routine treatment environment. In this chapter, we examined the feasibility and effectiveness of two evidence-based outpatient psychotherapeutic interventions for patients with alcohol use disorders. Furthermore we assessed whether pre-treatment patient characteristics are predictive of treatment outcomes for these interventions.

Chapter 5 describes the development and psychometric evaluation of the Questionnaire for Learning Organizations (QLO). The QLO is based on the theory of Learning Organizations of Senge. Psychometric properties (reliability and factor structure) of the QLO are assessed for a group of treatment professionals and a group of supporting employees in substance abuse treatment centers, in order to evaluate the quality of the questionnaire.

Chapter 6 investigates whether ROM plays a role in enhancing learning capacity in terms of Senge’s theory of Learning Organizations. In a quasi-experimental pre-post design, we investigate whether treatment professionals who engaged in the ROM system show an increase in learning potential.

This thesis ends with a general discussion and a summary, where the conclusions from all chapters are compiled and discussed and suggestions for future research will be given.