

What can and what cannot be concluded from patient-reported outcomes as a quality indicator in medical rehabilitation?

Beate Muschalla

Technische Universität Braunschweig, Department of Psychotherapy and Diagnostics

Abstract

Some weeks after inpatient rehabilitation treatment in Germany, patients are invited to fill in a questionnaire on their perception and satisfaction with the rehabilitation treatment. This procedure has been established due to the need of external quality assurance by the payers (e.g. the German Federal Pension Agency) according to social law SGB IX § 37. These patient reported outcomes and perceptions are included in the external quality assurance system by the German Federal Pension Agency. The question is: What do these observation data mean, what can be concluded from them in respect to quality of medical rehabilitation, and where are limitations?

Due to the heterogeneity of "quality of medical rehabilitation" (different rehabilitation settings, patient clientele, indications, profiles of clinics) there cannot be a global operationalization of a "good rehabilitation clinic"

Different instruments are necessary for different goals. It must be differentiated between observation for advertising on the one hand, and observation for benchmarking and controlling on the other hand.

When patient reported outcomes are lacking reliability and validity, they cannot be used for purpose of benchmarking and controlling. Patient reported outcomes may be used as a screening for special abnormalities. For example, the "last 5%" (those clinics with highest patients' dissatisfaction) could be visited or consulted and problem solving processes may be initiated: pension agency and clinicians might discuss the conditions of the respective clinic and if necessary find the problem and specific problem solutions. Judgments whether diagnostics, treatment and treatment coordination is correct (this is quality assurance in medical rehabilitation) can only be done by observation in the clinical routine. This observation should be done by visitations by specialists (physicians and therapists) of the respective medical indication.

Keywords

Quality assurance, quality management, controlling, benchmarking, patient reported outcomes, PRO

Nutzung von Rehabilitandenbefragungen im Rahmen externer Qualitätssicherung – Möglichkeiten und Grenzen unter Berücksichtigung der Datenqualität

Kurzfassung

Für die Rehabilitation besteht ein gesetzlicher Auftrag zur vergleichenden Qualitätssicherung durch die Rehabilitationsträger (§ 37 SGB IX). Die Deutsche Rentenversicherung hat Qualitätskriterien definiert, die in Rehakliniken flächendeckend zu diesem Zweck beobachtet werden. In diesem Beitrag wird ein zentrales Instrument dieser externen Qualitätssicherung – die Rehabilitandenbefragung – aus klinischmethodischer Sicht diskutiert.

Nach medizinischen Rehabilitationsbehandlungen (§ 42 SGB IX) werden routinemäßig zufällig ausgewählte Patienten von den Leistungsträgern angeschrieben und gebeten einen standardisierten Fragebogen zu ihrem Rehabilitationsaufenthalt auszufüllen. Hierbei wird erfragt, welche Behandlungen der Patient erfahren hat, mit welchen Behandlern er Kontakt hatte, ob Rehabilitationsziele subjektiv erreicht worden sind und wie der Patient den Aufenthalt insgesamt erlebt hat und bewertet. Methodisch handelt es sich hierbei um sogenannte Beobachtungsdaten. Es stellt sich die Frage, für welchen Zweck Daten aus den Rehabilitandenbefragungen im Hinblick auf die externe Qualitätssicherung medizinischer Rehabilitation verwendet werden können und wo die Grenzen der Verwendbarkeit liegen.

Aufgrund der Heterogenität dessen, was „Qualität medizinischer Rehabilitation“ im Einzelnen mitdefiniert (unterschiedliche Rehasettings, Indikationen, Patientenkontexte, Versorgungsprofile der Kliniken), kann es zum Zweck der Steuerung keine Globaloperationalisierung einer „guten Klinik“ geben.

Grundsätzlich sind für unterschiedliche Ziele und Beobachtungsaufgaben unterschiedliche Beobachtungsinstrumente erforderlich. Es ist im Rahmen der Qualitätssicherung der medizinischen Rehabilitation insbesondere zwischen Beobachtung zum Zwecke von Berichterstattung (Public Reporting) und Beobachtung zur (Belegungs)Steuerung von Kliniken zu unterscheiden.

Wenn Rehabilitandenbefragungsdaten nicht im Sinne einer Abbildung von Qualität medizinischer Rehabilitationsbehandlung (d.h. komplexer Diagnostik- und Behandlungsaktivitäten, vgl. § 42 SGB IX) valide sind, ist ihre unmittelbare Verwendung als Benchmarking- oder gar Steuerungsinstrument ausgeschlossen. Sie können jedoch genutzt werden als Screening für besondere Auffälligkeiten: Beispielsweise kann man sich die Kliniken mit erhöhtem Anteil an Unzufriedenen ansehen und im Rahmen von Beobachtung und Exploration vor Ort (Visitationen, Strukturierter Qualitätsdialog SQD) herausfinden, was das Problem ist. Aussagen darüber, ob in Kliniken „medizinische Rehabilitation“ (§ 42 SGB IX), d.h. Diagnostik und Behandlung(skoordination) für die jeweiligen individuellen Patienten, lege artis durchgeführt wird, sind nur im Rahmen von Visitationen (durch Therapeuten und Ärzte der jeweiligen Indikation) und direkter Beobachtung in der Klinikroutine möglich.

Schlüsselwörter

Qualitätssicherung, Qualitätsmanagement, Steuerung, Benchmarking, Rehabilitandenbefragung

1 Conditions of quality assurance in German rehabilitation clinics

Some weeks after an inpatient rehabilitation treatment, a random selection of patients are invited by the German Federal Pension Agency to fill in a standardized questionnaire on their satisfaction and outcome of rehabilitation. Patients are asked which treatments they received, which professionals conducted the treatments, whether and to which amount rehabilitation goals have been reached and how the patient liked the rehabilitation stay as such.

The results of this observation are used within the quality assurance program of the German Person Agency as “patient reported outcomes” (Bassler et al., 2007; DRV, 2017a; Klosterhuis, 2010). Presently the following quality indicators are used within the extern quality assurance program (DRV, 2017a):

- treatment satisfaction (patient reported outcome)
- subjective treatment success (patient reported outcome)
- peer-review of medical reports (Farin et al., 2003; Strahl et al., 2016)
- therapeutic care (catalogue of therapeutic services)
- following rehabilitation treatment standards (Farin et al., 2018)

They are decisive for the judgment of the quality of a rehabilitation clinic by the pension agency, and also decisive for “interventions” in case of lack of quality (e.g. structured quality dialogue, DRV, 2017b).

Legal basis for quality assurance is the German Social law (§ 37 SGB IX) which explicitly demands for comparative quality analysis as a

basis for effective quality management. Similar systems exist in other countries and generally in the health care system. They are (critically) discussed (Köhn et al., 2016; Saver et al., 2015; Kolkman et al., 2004).

In this article, „rehabilitation quality“ will be understood as the quality of lege artis conducted diagnostics, medical treatment, and treatment coordination. This is what medical rehabilitation in core means. Official patient information given by the German Federal Pension Agency defines it as follows „Medical rehabilitation focuses the treatment of health- and functioning disorders“ (DRV, 2017c). In each individual case all rehabilitation processes require medical decision making by specialists of the respective indication. This is according to the general rules of professional conduct in medicine and psychotherapy (Deutscher Ärztetag, 2018; BPtK, 2007). According to the biopsychosocial medicine model (ICF, WHO, 2001) treatment can focus on symptom reduction, capacity training and context-oriented treatments such as assistance systems, or coordination activities such as initiating a stepped reintegration at work. There is a wide variety of treatment options for rehabilitative purposes in inpatient and outpatient treatment (TOPPP Checklist, Linden et al., 2018).

Also in German social law, the complexity and diversity of medical rehabilitation has been differentiated and listed (§ 42 SGB IX, Table 1). This makes clear that the declared main aim of medical rehabilitation is to avoid chronification of illness, and promote patients’ participation in general and professional life. This is realized by means of individual diagnostic and complex individual treatment according to medical standards.

<p>Medical rehabilitation comprises especially</p> <ol style="list-style-type: none"> 1. treatment by physicians and other health care professions, including treatments for activating self-help 2. early detection and early support for children with disabilities 3. pharmaceutical products and surgical materials 4. remedies including physicalic, speech and occupational therapy 5. psychological or medical psychotherapy 6. assistive devices 7. vocational and stress tests <p>Services may be medical, psychological and educational aids, which must be chosen individually in each case</p> <ol style="list-style-type: none"> 1. support in coping with illness 2. support for self-help 3. information and counselling for partners, or relatives, or supervisors and workmates of the patients (patient's consent provided) 4. relaying of contacts to self help institutions or counselling 5. trainings of social competency and interaction capacities, as well as coping with critical life events, 6. training of capacities of daily living 7. motivation and guidance in choosing rehabilitation and support means

Table 1

German Social Law Book IX (§ 42 SGB IX): Treatment methods within medical rehabilitation

2 Question: What can be learned and derived from the patients' questionnaire data (patient reported outcomes)?

Under the above introduced conditions, the question arises how to interpret the data from these patients' questionnaires, and how they can be used within the frame of quality assurance. Where are limitations of interpretation and conclusions? In the following I will give some basic remarks on this question from a methodological point of view. Some general information on data formats will precede (Bortz, 1999; Bortz & Döring, 2003; Field, 2012).

3 Data formats: Experimental studies vs. observation studies

A basic differentiation must be made between

- data from experimental studies (i.e. from systematic variation of an independent variable and analysis of a dependent variable, e.g. within a randomized controlled study) which allows causal conclusion and
- data from observation studies, which do not allow causal assumptions, but serve different aims, depending on their type.

3.1 Types and aims of observation studies

Different types of observational studies can be distinguished (Table1): Free observation is the casuistic observation of individual cases. It can be done by unstructured or criteria-guided observation and corresponds to the approach of qualitative research. The aim is to gain insights in processes or phenomena, or satisfying (scientific) curiosity. Quantifying observation generates frequencies and is based on observation according to a catalogue of phenomena. Quantitative comparisons compare frequencies which occur in different settings. According to a catalogue of phenomena, observations are conducted in different defined settings. Qualifying observation sets the additional criterion whether thresholds are reached. We speak of qualitative comparisons (benchmarking) when comparative observation is done by considering thresholds. Within the frame of hypothesis generating variance enlightenment, correlations and variance in multidimensional data sets are investigated. The data do not allow causal interpretation. Finally, there is observation with the aim of controlling systems and processes, i.e. observation followed by an intervention.

Tabelle 2
Types of observation
and examples

Type of observation	Fictive example	Fictive example rehabilitation clinic
Free observation	Observing what deer are doing on a meadow	Observing what psychologists do in a rehabilitation clinic
Quantifying observation	How many deer are there in Brandenburg?	How many psychologists are employed in a rehabilitation clinic?
Quantitative comparisons	How many deer are there in Brandenburg and how many in the Saarland?	How many psychologists are there in pension agency owned rehabilitation clinic and in private rehabilitation clinics?
Qualifying observations	Are there too many deer in Brandenburg?	Are there too many psychologists in pension agency owned rehabilitation clinics?
Qualitative comparisons (Benchmarking)	There are too many deer in the Saarland and too little in Brandenburg (even though there might be more deer in Brandenburg than in the Saarland as counted in absolute frequencies)	There are too many psychologists in person agency owned rehabilitation clinics but too little in private rehabilitation clinics.
Hypothesis generating variance enlightenment	There is a significant correlation between storks and deer in Brandenburg	There is a correlation between the number of psychologists and physicians in rehabilitation clinics.
Observation for controlling of systems and processes	Deer must not enter fields.	Psychologists are not allowed to give medication.

4 Methodological conditions

4.1 Reliability, Validity, Utility, Reactivity

Up from the quantifying observation, categories for observations must be available. For these, data for reliability, validity, and utility must be available.

Reliability and objectivity mean that two observations come to the same result. Reliability can vary between 0 (no agreement) and 1 (complete agreement). Validity means that the observation categories do indeed observe the object of interest (e.g. deer and not roe deer, or only psychological psychotherapists and no diploma/master psychologists) and nothing else (Buntin et al., 2017). Utility is the relation of specificity (correct positive/healthy) and sensitivity (correct negative/ill) in dependence of the epidemiological distribution in the observation setting. For example, in case there are only deer but no roe deer in the Saarland, then utility is (considering the same specificity and sensitivity) different than in Brandenburg, where we find almost only roe deer but not deer. In case there are 99 deer in the wood and only one roe deer, and the specificity of diagnostic of deer is high (95% correctly classified deer), there will be 5 deer wrongly classified as roe deer and only one correct as roe deer (sensitivity). This means that 5 out of the 6 as roe deer classified animals are in reality deer. The same is valid for

questions like whether psychologists do behavior therapy dependent on how many behavior therapists and how many psychoanalysts are in the clinic.

Reactivity means the degree of change which an observation causes in the object of observation. Whenever the ranger arrives, the deer disappears, which results in a smaller number of deer counted. When the visitor arrives in a clinic all part-time employees are present.

4.2 Standards for reliability and validity depending on the observation aim

In free observations reliability is regularly low, but instead certain validity can be assumed. In quantifying observations and quantitative comparisons, moderate test's quality criteria values of 0.6 are sufficient.

In qualifying observations, qualitative comparisons, and hypothesis generating variance enlightenment the requirements for test's quality criteria depend on the number of observations. The smaller the number of observations, the higher the test quality must be.

In observations for controlling and managing system and processes the test quality values must be near 1. For example, a body weight scale which a person uses occasionally for measuring his weight, must show a certain range, a medical scale in contrast which is needed to exactly control medication in case

of water retention, must show hundred-percent exact values of weight.

5 Use of the different observation methods

For reaching different aims and observation tasks, different observation means are necessary. Especially observations aiming reporting on the one hand, and observations for controlling on the other hand have to be distinguished.

5.1 Presentation of rehabilitation and reporting/advertising statement

In case it is necessary to show that nature protection assures survival of deer or that rehabilitation leads to a good outcome, then qualifying observations are necessary. They require quantification as well as a threshold for quality. This result depends on the investigator and not the object of observation. It is the investigator who decides

- what exactly shall be investigated (therapists, diploma/master psychologists, psychological psychotherapists)
- how the questions are posed
- which thresholds of quality shall be chosen.

Depending on the aim what shall be reported, it can be presented that an item of observation is bad, or good: In the first case, high thresholds can be chosen and questions formulated so that a left-leaning distribution arises. In contrast, on basis of the same basis one could say that rehabilitation is in 99% of cases good, by choosing a low threshold and thereby initiating a right-leaning distribution. There is no truth, but target margins which have been defined by the interest carriers.

For credibility purposes it may be ideal when more than 70% of the respondents appear to be very good, 25% good and 5% bad. It can then be explained that everything is alright and the “bad” 5% will be consulted to find out what the problem is if any, and eventually find problem solutions. In the context of rehabilitation quality, the bad clinics can be consulted by means of visitations and the structured quality dialogue (DRV, 2017a).

5.2 Qualitative comparisons (Benchmarking)

In case it is required to make a ranking, i.e. saying whether Brandenburg or Saarland is better in caring for deer, or whether state or private rehabilitation clinics are better, qualitative comparisons are necessary.

This requires a power calculation, i.e. it must be ensured that differences (given reliability of the instrument) become visible. This in turn requires sufficient validity (deer and roe deer must not be confused), and the latter in turn requires non-reactivity (deer must not flight in reaction to the observation).

When not only a qualitative comparison (where are more deer) is intended, but an evaluation and possibly rating of nature protection, additional intervening variables must be considered (in Brandenburg there are many deciduous woods, and in the Saarland many coniferous forests). Basically it must be kept in mind that there can never be a confident comparison even in case of using multivariate methods (e.g. multivariate conditions of degree A in school reports in Brandenburg or in the Saarland). To have a valid benchmarking, large safety margins are required, i.e. by using a significance level of $p < .001$, and effect size differences of at least 0.5.

5.3 Controlling

In case observation shall be used for controlling purposes (e.g. flight attitude in flight from Rio to Paris), then extremely high requirements have to be fulfilled:

- A reliability of 1 is needed.
- Data must have an absolute validity.
- Data must be available on time.
- Data must allow concrete conclusions.
- The derived interventions must be effective, i.e. must lead to a positive change.
- It must be shown that the chosen change meets the intended goal (problem of surrogate parameters, Saver et al., 2015).

Standards which are presently sufficient for scientific purposes in the domain of psychological characteristics¹ are not sufficient when processes shall be controlled which directly affect medical professional actions. The reason lies in the insufficient utility respective predictive validity: For a psychometrically established instrument with a sensitivity of about 0.8 and a specificity of 0.7 we find that in a sample of 100 patients with a true prevalence

¹ An example are psychometric developments of patient self report scales based on patient self report data (Schmidt et al., 2018; Spyra et al., 2006; Kriz et al., 2008)

of 25%, there will be more patients wrongly classified as ill (25) than correctly identified as ill (20) (Lenhard & Lenhard, 2014). Due to this fact, much more complex requirements are imposed on clinical diagnostics as compared with research diagnostics (Linden & Muschalla, 2012)².

6 Conclusions for the extern quality assurance of rehabilitation clinics

6.1 Possible problems with patient reported outcomes

Possibilities and limitation for the use of patient reported outcome have been discussed for many years (Bassler et al., 2007; Klosterhuis, 2010; Farin & Jäckel, 2011; Widera, 2010; Widera et al., 2011; Linden, 2011). By using patient reported outcomes for quality assurance the problem arises that observation, benchmarking, and controlling are mixed up. By this, none of the different aims can be reached.

Until now it is unclear how to deal with conflicts of objectives: Is the aim to gain high patient satisfaction, or is the aim to have valid socio-medical judgments? It is unclear what can be concluded from patient reported outcomes, because of missing validity. The validity problem concerns different aspects of the patient questionnaire, i.e. quality of medical care, which therapists have been involved, and rehabilitation result.

When results from patient investigation and other quality indicators are feedbacked to the clinic only after a long time of several months, they are not useful for decisions on interventions, because nobody knows what the present situation is.

From the patients investigation as such no concrete decisions for actions or interventions can be derived.

In order to detect reasons for dissatisfied patient judgments (within the patient reported outcomes) the clinic (colleagues) has (have) to be visited. Within the frame of extern quality assurance the pension agency reports the re-

sults from the patient investigation to the clinics. The clinicians may use these information for adjusting their intern quality management (Nübling et al., 2018).

As a reaction towards criticism on the benchmarking based on patient reported outcomes, a statistical adjustment procedure was implemented into the extern quality assurance process. Adjustment by means of multivariate analysis can give hints about variance bond by third variables and shows the correlation pattern dependent on the variables included in the analysis. However, adjustment cannot fully control the multivariate heterogeneity of the observation objects. Multivariate analyses are dependent on which variables are included in the analysis and how valid these variables are. For example, in comparison of clinics in the city or in the country side, a wide variety of "city" is possible (e.g. Berlin versus Teltow). When important variables are not considered, misinterpretations may occur³. Furthermore, the general differences of patient groups and clinic specializations must be considered when judging quality of clinics. For example, within the frame of internal medicine, there are clinics which treat mainly illnesses with positive prognosis (e.g. diabetes) and other clinics treat rather patients with negative prognosis (e.g. inflammatory bowel diseases). Another example is the heterogeneous frequencies of work-reintegration problems in patients with different diagnosis in the indication group gastroenterology (Streibelt et al., 2017).

These examples show that in each case empirical observation of the clinics' specialties is necessary in order to understand concrete differences in patient reported outcomes. Empirical observations are necessary for understanding which variables are relevant for the explanation of differences. In sum, the multivariate heterogeneity of the objects of observation cannot be completely controlled, despite trials for multivariate adjustment.

When there cannot be reliable conclusions from the patient investigations, there cannot be wide reaching action: clinics cannot be controlled on basis of these data. Clinics cannot be compared against each other and no

² An example for the necessary high quality standards in rehabilitation medicine action in comparison to research is the diagnostic of mental health problems: For research purposes (e.g. epidemiological data on the distribution of mental health problems in 1000 patients) a standardized research interview or a questionnaire is sufficient. Rehabilitation medicine diagnostics in each single case however requires a complex somatic and psychopathological investigation, exploration of the medical history including treatment trials and treatment course (Linden & Muschalla, 2012). Clinical diagnostic cannot be based on algorithmic counting of patient reported symptoms (Linden & Rath, 2014). Clinical diagnostic in medical rehabilitation requires valid structure quality, i.e. qualification of the professionals, as well as adequate time resources).

³ For example, in a comparison of patients in inpatient treatment and patients in day clinic setting, patients in day clinic setting showed worse improvement. When including the type of intake (freewillingly versus sent by insurance company) and severity of illness, then it became clear that not the setting (day clinic or inpatient) was associated with the improvement, but the type of intake (Geiselmann & Linden, 2001).

political conclusions and decisions can be based on these data.

In case invalid conclusions are drawn from the data (comparisons of clinics, controlling and interventions), then side effects may occur: Clinicians may come into self-sufficiency or demotivation. Clinics may feel authorized to use these data for advertising, without a valid basis.

When a supervisory authority implements instruments for controlling, then they take responsibility for the resulting effects in respect to political, juristic, ethical perspectives and public reception. Highest professional standards have to be set for such controlling processes: Quality standards which are sufficient for scientific investigations (e.g. scale developments based on patient self-report measures, Schmidt et al., 2018; Sypra et al., 2006; Kriz et al., 2008), are however insufficient when processes shall be controlled which have immediate effects on professional actions, i.e. what clinicians do with patients in rehabilitation clinics.

The reason lies in reliability, specificity and validity and utility respective the predictive value, as explained above.

There is until now no evidence which could show consistently that patients' self reports on structure-, process-, and outcome quality aspects of medical rehabilitation are highly correlated with medical personal's reports of the same aspects of investigation. There are findings that patients' subjective judgments are correlated with their later professional integration, (Nübling et al., 2018). However, these are longitudinal observation data and do not allow the causal conclusion that a subjectively good rehabilitation success is caused by good rehabilitation treatment. Patients with worse subjective success had a worse status even before rehabilitation.

The German Federal Pension Agency is challenged to fit their external quality assurance system to divergent aims: towards politics the comparing report of rehabilitation quality (§ 37 SGB IX), towards the clinics the aim of a highly valid description of medical rehabilitation quality. As a supervisory authority, they are responsible for quality of medical rehabilitation. They are also politically and in public responsible for effects resulting from certain controlling procedures they implement.

6.2 Solutions

It can be made precise which information are necessary for which aim and which wider reaching aims are addressed. It can then be cleared from a scientific perspective what

can be done. A global operationalization of a "good clinic" is not possible. As a solution, a differentiation can be made between public reporting (advertising) on the one hand, and benchmarking or controlling on the other hand. These two aims need different methods.

1. **In public it can be reported** that rehabilitation is a health care sector of good quality. This is an advertising task (see 5.1) and can be fulfilled accordingly, e.g. by using patient reported outcomes. Thereby no comparison of single clinics should be made (benchmarking). Instead, the percentages of very good and good results (in sum 95%, e.g. Giesler et al., 2017: 90% fulfill therapy standards) should be reported, and the note, that the problematic 5% (5th percentile) will be regarded individually by means of visitation or structured quality dialogue. When **patient reported outcomes** are not valid for quality of medical rehabilitation, then their use for benchmarking und controlling is not possible. But, they could be used as a **screening instrument** for detecting the 5-10% of clinics with higher percentages of dissatisfied patients. Then these clinics can be visited and thereby observed and explored if there is a quality problem and eventually which problem. This procedure concerning the "last 5%" can be done similarly with peer-review and other quality criteria (Klosterhuis, 2010).
2. Finally, the question whether treatment and treatment coordination are conducted correctly in rehabilitation clinics (i.e. **quality assurance of medical rehabilitation**) can only be answered by means of **visitations in the clinic routine**. Hereby specialists (physicians and therapists) of the respective indications are of importance in order to evaluate the indication-specific medical diagnostics and treatment in a valid manner (Heudorf et al., 2015; Meixner et al., 2006; Neuderth et al., 2017; Salzwedel et al., 2018). The new instrument of the "structured quality dialogue" by the German Pension Agency (Ostholt-Corsten & Weinbrenner, 2018) is an instrument which allows an individual problem detection and problem solving in case the "quality data" of a clinic are poor. Thus, the controlling does not automatically induce any action (e.g. occupancy reduction) but gives way for qualitative investigation and judgment by experts.

7 Limitations and summary

In this paper one instrument within the external quality assurance system in medical rehabilitation has been discussed: patient reported outcomes on treatment satisfaction and subjective treatment success. The other quality criteria could be discussed similarly.

The author of this paper defines rehabilitation quality according to the contents of medical rehabilitation and according to the legal definition (German social law, § 42 SGB IX): complex and coordinated medical diagnostic and treatment of chronic illnesses, with focus on life and work participation. The German Federal Pension Agency chose an operational definition of quality according to structure, process and result quality and hereby used existing parameters (patient reported outcomes, catalogue of therapeutic services, rehabilitation treatment standards). There will be the question of relevance and appropriateness of these parameters and their relative weight within the quality assurance processes, e.g. in the present pilot project for choice of clinics (DRV, 2017d).

Presently, as seen from methodological and clinical perspective, the instrument of the structured quality dialogue is an adequate means for clearing possible “quality problems” in certain clinics. The dialogue allows pension agency and clinicians to live observe the clinic and discuss its specialties and whether and eventually which “quality problems” may be present and how they can be solved. The patient reported outcomes can only function as a screening for detecting possible problems. However, patient reported outcomes cannot be the basis for controlling processes and decisions for occupancy rates in a clinic without observing the clinic.

References

- Bassler, M., Nosper, M., Follert, P., Böwering, L., & Polak, U. (2007). Datenquellen für eine kontinuierliche Qualitätsverbesserung in der medizinischen Rehabilitation. *Die Rehabilitation*, 46, 155–163.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J., & Döring, N. (2003). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- BPTK. (2018). *Muster-Berufsordnung für die Psychologischen Psychotherapeutinnen und Psychotherapeuten und Kinder- und Jugendlichenpsychotherapeutinnen und Kinder- und Jugendlichenpsychotherapeuten*. Dortmund: Deutscher Psychotherapeutentag.
- Buntins, M., Buntins, K., & Eggert, F. (2017). Clarifying the concept of validity: From measurement to every day language. *Theory & Psychology*, 27, 703–710.
- Deutscher Ärztetag. (2018). *Muster-Berufsordnung für die in Deutschland tätigen Ärztinnen und Ärzte*. *Deutsches Ärzteblatt*, 05/2018, A1–A9.
- DRV. (2017a). *Qualitätssicherung in der medizinischen und beruflichen Rehabilitation*. Berlin: Deutsche Rentenversicherung.
- DRV. (2017b). *Der Strukturierte Qualitätsdialog mit Reha-Einrichtungen als ein Instrument der Reha-Qualitätssicherung der Deutschen Rentenversicherung*. Berlin: Deutsche Rentenversicherung.
- DRV. (2017c). *Medizinische Rehabilitation: Wie sie Ihnen hilft*. Berlin: Deutsche Rentenversicherung.
- DRV. (2017d). *Beschaffung von Leistungen zur medizinischen Rehabilitation*. SGB VI § 138 Abs. 1 Satz 2 Nr. 4 Buchstabe a und Abs. 2 Satz 1. Verbindliche Entscheidungen der Deutschen Rentenversicherung Bund, Inkrafttreten 2017. RV aktuell, 2017. Online abgerufen am 30.07.2018 unter https://www.deutsche-rentenversicherung.de/Bund/de/Inhalt/6_Wir_ueber_uns/01_infos_zum_unternehmen/03_unternehmensprofil/012_selbstverwaltung/03_verbindliche_entscheidungen/2017/20170825_leistungen_zur_medizinischen_reha.html
- Farin, E., Carl, C., Lichtenberg, S., Jäckel, W.H., Maier-Riehle, B., & Rütten-Köppel, E. (2003). Die Bewertung des Rehabilitationsprozesses mittels des Peer-Review-Verfahrens: Methodische Prüfung und Ergebnisse der Erhebungsrunde 2000/2001 in den somatischen Indikationsbereichen. *Die Rehabilitation*, 42, 323–334.
- Farin, E., Wirth, A., Quaschnig, K., Meyer, T., Nowik, D., Block, I., Vogel, H. & Volke, E. (2018). Die Aktualisierung der Reha-Therapiestandards der Deutschen Rentenversicherung. *Die Rehabilitation*, 57, 48–54. doi: 10.1055/s-0043-124309.
- Farin, E., & Jäckel, W.H. (2011). *Qualitätssicherung und Qualitätsmanagement in der medizinischen*

- Rehabilitation. *Bundesgesundheitsblatt*, 54, 176–184.
- Field, A. (2012). *Discovering Statistics using SPSS*. London: Sage.
- Geiselman, B., & Linden, M. (2001). Vollstationäre, tagesklinische und kombiniert stationär-teilstationäre psychosomatische Rehabilitation im Vergleich. *Verhaltenstherapie und Verhaltensmedizin*, 22, 433–451.
- Giesler, J.M., Faust, T., Ernst, J., Kuhnt, S., Mehnert, A., & Weis, J. (2017). Ausgewählte Merkmale der Struktur- und Prozessqualität der psychosozialen Versorgung in ambulanten psychosozialen Krebsberatungsstellen in Deutschland. Eine Bestandsaufnahme. *Gesundheitswesen*, 08-2017. doi: 10.1055/s-0043-111233.
- Heudorf, U., Färber, D., Mischler, D., Schade, M., Zinn, C., Nillius, D., & Herrmann, M. (2015). Multiresistente Erreger in Rehabilitationseinrichtungen im Rhein-Main-Gebiet, Deutschland, 2014: II. Ärztliche Risikoanalyse und Hygienemaßnahmen. *Die Rehabilitation*, 54, 375–381. doi: 10.1055/s-0035-1564099.
- Klosterhuis, H. (2010). Reha-Qualitätssicherung der Rentenversicherung – eine kritische Bestandsaufnahme. *RV aktuell*, 08/2010, 260–268.
- Köhn, S., Schlumbohm, A., Brünger, M., Menzi, L., Schwarzenbrach, S. Busch, P., & Spyra, K. (2016). Externe Qualitätssicherung in der medizinischen Rehabilitation in der Schweiz. Klinikvergleichen- de Ergebnisqualität in der muskuloskelettalen Rehabilitation. *DRV-Schriften*, 109, 103–105.
- Kolkman, F.W., Vilmar, K., & Stobrawa, F. (2004). Entprofessionalisierung ärztlicher Berufsausübung. *Deutsches Ärzteblatt*, 101, 1409–1414.
- Kriz, D., Nübling, R., Steffanowski, A., Wittmann, W.W., & Schmidt, J. (2008). Patientenzufriedenheit in der stationären Rehabilitation: Psychometrische Reanalyse des ZUF-8 auf der Basis multizentrischer Stichproben verschiedener Indikationen. *Zeitschrift für Medizinische Psychologie*, 2-3/2008, 67–79.
- Lenhard, W. & Lenhard, A. (2014). Berechnung von Testgütekriterien für Screeningverfahren. verfügbar unter: <http://www.psychometrica.de/testkennwerte.html>. Bibergau: Psychometrica.
- Linden, M. (2011). Möglichkeiten und Grenzen der Beobachtungsforschung in der psychosomatischen Rehabilitation. Vortrag im Rahmen des Forschungskolloquiums der Forschungsgruppe Psychosomatische Rehabilitation. Rehazentrum Seehof, Teltow.
- Linden, M., & Muschalla, B. (2012). Standardized diagnostic interviews, criteria, and algorithms for mental disorders: garbage in, garbage out. *European Archives of Psychiatry and Clinical Neuroscience*, 262, 535–544.
- Linden, M., Muschalla, B., Noack, N., Heintze, C., & Döpfmer, S. (2018). Treatment Changes in General Practice Patients with Chronic Mental Disorders following a Psychiatric-Psychosomatic Consultation. *Health Services Research and Managerial Epidemiology*, 5, 1–6.
- Linden, M., & Rath, K. (2014). The impact of the intensity of single symptoms on the diagnosis and prevalence of major depression. *Comprehensive Psychiatry*, 55, 1567–1571.
- Meixner, K., Lubenow, B., Brückner, U., & Gerdes, N. (2006). Weiterentwicklung und Validierung eines Verfahrens zur Visitation von Rehabilitationseinrichtungen. *Die Rehabilitation*, 45, 152–160.
- Neuderth, S., Saupe-Heide, M., Brückner, U., Gross, B., Wenderaath, N., & Vogel, H. (2012). Wissenschaftliche Begleitung der Weiterentwicklung des Visitationsverfahrens für Rehakliniken der Deutschen Rentenversicherung Bund. Das Projekt „Visit II“. *Die Rehabilitation*, 51, 171–180. doi: 10.1055/s-0031-1280804.
- Nübling, R., Kaiser, U., Kriz, D., & Schmidt, J. (2018). Bedeutung von Patient Reported Outcomes (PROs) für Qualitätssicherung/Qualitätsmanagement in der medizinischen Rehabilitation - Fortlaufendes Qualitätsmonitoring durch kontinuierliche Rehabilitandenbefragungen. *Praxis Klinische Verhaltensmedizin und Rehabilitation*, 102, xx-xx.
- Salzwedel, A., Haubold, K., Barnack, B., Reibis, R., & Völler, H. (2018). Indikatoren der Ergebnisqualität kardiologischer Rehabilitation. Ergebnisse einer Delphi-Befragung von Mitgliedern der Deutschen Gesellschaft für Prävention und Rehabilitation von Herz-Kreislauferkrankungen e.V. (DGPR). *Die Rehabilitation*, ahead of print, doi: 10.1055/s-0044-101048.
- Saver, B.G., Martin, S.A., Adler, R.N., Candib, L.M., Deligiannidis, K.E., Golding, J., Mullin, D.J., Roberts, M., & Topolski, S. (2015). Care that Matters: Quality Measurement and Health Care. *PloSOne*, 12(11): e1001902. doi: 10.1371/journal.pmed.1001902
- Schmidt, J., Nübling, R., Kriz, D., Kaiser, U., & Wittmann, W.W. (2018). Die Kurzskala BESS zur direkten Veränderungsmessung des gesundheitlichen Befindens. Psychometrische Überprüfungen für den Bereich der Medizinischen Rehabilitation. *GfQG-Schriften* 1.
- Spyra, K., Erhart, M., Müller-Farnow, W., & Rieger, J. (2006). Das KlinikSpezifischePatientenInventar (KSPI) zur Beurteilung der Leistungen in den somatischen und psychosomatischen Sucht-Rehabilitation (KSPI-SO/PS). *Praxis Klinische Verhaltensmedizin und Rehabilitation*, 2006, 71, 47–65.
- Strahl, A., Gerlich, C., Wolf H.-D., Gehrke, J., Müller-Garnn, A., & Vogel, H. (2016). Qualitätssicherung in der sozialmedizinischen Begutachtung durch Peer-review – ein Pilotprojekt der Deutschen Rentenversicherung. *Gesundheitswesen*, 78, 156–160. doi <http://dx.doi.org/10.1055/s-0034-1390450>

- Streibelt, M., Bethge, M., Gross, T., Herrmann, K., Ustaoglu, F., & Reichel, C. (2017). Predictive Validity of a Screening Instrument for the Risk of Non-Return to Work in Patients with Internal Diseases. *Archives of Physical Medicine and Rehabilitation*, 98, 989–996. doi: 10.1016/j.apmr.2016.12.022
- WHO. (2001). *International Classification of Functioning, Disability and Health ICF*. Geneva: World Health Organisation.
- Widera, T. (2010). Aktuelles aus der Reha-Qualitätssicherung- Neue Ergebnisse der Rehabilitandenbefragung. *RV aktuell*, 4, 2010, 153–159.
- Widera, T., Beckmann, U., Zander, J., & Radoschewski, M. (2011). Wer hat Recht - Rehabilitand oder KTL-Dokumentation? *RV aktuell* 10, 2011, 291–298.

Prof. Dr. Beate Muschalla

*State-licensed behavior therapist
and supervisor*

*Technische Universität Braunschweig,
Department Psychotherapy and Diagnostics
Humboldtstraße 33
38106 Braunschweig
b.muschalla@tu-bs.de*