Электронный журнал «Психологическая наука и образование»
www.psyedu.ru / ISSN: 2074-5885 / E-mail: psyedu@mgppu.ru 2012, №4

Evaluation in Rehabilitation: Outcomes, Assessments, and Measurement of Change

J. Bengel,
Albert-Ludwigs-University of Freiburg, Germany

M. Kraft,
Albert-Ludwigs-University of Freiburg, Germany

O. Mittag,
University Medical Center of Freiburg, Germany

The evaluation of rehabilitation interventions is of essential significance in rehabilitation and rehabilitation research. In this context, the present article deals with the most important aspects concerning the measurement of success and change. Pivotal outcome criteria and indicators as well as the most often applied self report instruments are introduced. The IRES-3 (Indicators of Rehabilitation Status), developed in Germany, is described in more detail. Following questions regarding the measurement of change, alternative assessment methods, namely goal-oriented measure and adaptive testing based on Item-Response-Theory, are discussed.

Key words: Diagnostic measures, assessment methods, outcome evaluation, quality of life, item response theory, rehabilitation.

Introduction

In rehabilitation research, measuring success is of particular importance, especially considering the aspects of legitimization and further development of rehabilitative interventions.

Essential are the definition, the operationalization, and the measurement of relevant outcome parameters. Outcome research in rehabilitation goes beyond disease symptoms and impairments of functioning. The 'International Classification of Functioning, Disability and Health' (ICF; DIMDI, 2004; Peterson, 2005) developed by the 'World Health Organization' (WHO) serves as the framework for determining outcome criteria.

1. ICF as a conceptual framework

The ICF is a classification system that allows for the classification of health-related problems on a somatic level and their consequences regarding the ability to perform in the realms of work and everyday life as well as social participation (Bruyere, Van Looy, & Peterson, 2005). The

multifactorial model attempts to take into consideration the entire life situation of an individual. Generally, constraints concerning state of health pertain to the three central dimensions: The (1) impairment of body structures and body functions entails (2) activity limitations by which (3) the possibilities to participate in social and societal life are restrained. Disability is regarded as the result of an interaction of health problem and contextual factors, i.e. environmental factors (e.g. prejudices, barriers) and personal factors (e.g. resources, coping strategies). Body structures are parts of the body (organs, limbs and their components), body functions are elementary physiological and psychological functions. Activity stands for everyday activities (e.g. eating, clothing, personal hygiene); constraints are described as "activity limitations". Participation describes the extent to which a person is involved in everyday life in a societal context and the extent to which the individual is able to fulfill roles (e.g. work-wise, in the family); constraints are denoted as 'participation restrictions'. Thus, the ICF reflects the entirety of all possible constraints that are relevant success criteria for rehabilitative interventions.

2. Criteria and indicators of rehabilitative success

The question of rehabilitative success requires consensus regarding the relevant success criteria. These can be organized as follows (see also Biefang, Birkner, Thien, Härtel, & Bullinger, 1997; Bullinger & Ravens-Sieberer, 2000):

- psychosocial indicators (e.g. coping with disease, quality of life, independence in everyday life, depression, treatment satisfaction)
- biomedical indicators (e.g. ability to move, blood glucose level)
- economical indicators (e.g. costs, earning ability)

Besides organizing success criteria in terms of content, several stakeholders can be distinguished who have different expectations for rehabilitation: patients, therapists/personnel, sponsors, institutions, experts, society.

For the psychosocial indicators, subjective measures such as subjective health and health-related quality of life are frequently used. Quality of life is conceptualized as a multidimensional construct referring to well-being and ability to function, either from the perspective of the respective individual or from the perspective of external observers (Bullinger, 2000). Here, quality of life can be perceived as a comprehensive concept (e.g. general life satisfaction) or domain specific (e.g. quality of work-life). Individual-centered approaches for assessing quality of life pertain to the difference between aspired and subjectively attained goals. In dimensional approaches, quality of life is conceived as interindividual differences in physical, psychological, social, functional, material or spiritual aspects. Health economic approaches define quality of life by means of the preference for different health-relevant scenarios representing starting points for cost-benefit-analyses (QALY:

Costs per Quality Adjusted Life Year) (Freund & Ziegelmann, 2009). The subjective weighting of individual indicators can change in the course of rehabilitation (Bernhard, Lowy, Mathys, Herrmann, & Hürny, 2004).

3. Multidimensional self-report instruments

Particularly in the domain of rehabilitation, the inclusion of patients' self-assessments is important as one's own perception of e.g. health status, performance potential and motivation has crucial effects on coping with disease and on actual behavior. Despite similar physical impairments, subjective perception, appraisal, and processing of the consequences of disease and disability vary considerably. There also is variation regarding the subjective perception, appraisal, and processing of treatment-induced changes in individual symptom areas.

Due to the heterogeneous indication groups and different clinical pictures, generic and disease-specific instruments have to be distinguished. Generic instruments (i.e. instruments applying to different diseases/indications) allow the assessment of aspects which reflect the patient's strains or constraints independently of the indication. Generic instruments also allow the assessment of further characteristics important to the treatment process and to the evaluation of treatment success. Generic instruments constitute an important basis for an indication-independent comparison of the need and the quality of health care in rehabilitation. They also allow the identification of indication-independent relations (e.g. predictive value of treatment motivation for the improvement of health-related quality of life) and modes of action important to the outcome quality. Examples are:

FIM – Functional Independence Measure (Keith, Granger, Hamilton, & Sherwin, 1987)

IMET – Index for the Assessment of Health Impairments (Deck, Mittag, Hüppe, Muche-Borowski, & Raspe, 2007)

IRES-3 – Indicators of Rehabilitation Status (Bührlen, Gerdes, & Jäckel, 2005)

NHP – Nottingham Health Profile (Hunt, McKenna, McEwen, Williams, & Papp, 1981)

SF-36/SF-12 – Health Survey (Ware & Sherbourne, 1992 / Ware, Kosinski, & Keller, 1996)

WHOQOL-100 –WHO-Quality of Life Questionnaire (Szabo, 1996)

Exemplarily for these instruments, the IRES patient questionnaire (Indicators of Rehabilitation Status, 3. version; Bührlen et al., 2005; short version: IRES-24; Wirtz, Farin, Bengel, Jäckel, Hämmerer, & Gerdes, 2005) shall be described here. The IRES questionnaire comprises 144 items and was developed specifically for application in rehabilitation. Therein, the 'coping' of the rehabilitating individual is conceived to be a pivotal determinant of the rehabilitation process. In the IRES patient questionnaire, health related impairments are assessed by the dimensions

'physical health' and 'pain'. In order to adequately represent 'physical health', indicators for the sub-aspects 'symptoms heart/circulation' and 'symptoms musculoskeletal system' as well as the 'self-rating – health' and the 'satisfaction regarding health' are assessed. Further, in order to be able to document factors influencing the health impairments, aspects concerning the 'information status regarding the disease', the 'evaluation of information status', and the 'implementation of knowledge' are assessed via the dimension 'health-related behavior'. Psychosocial aspects are primarily presumed to be decisive determining factors for the pivotal variable coping process. In the realm of psychological determinants of the coping process (IRES dimension 'mental health'), particularly 'depression', 'anxiety', 'vital exhaustion', 'sense of self-worth' as well as 'cognitive performance ability' are assessed. 'Social support', ‚satisfaction regarding family', as well as ‚satisfaction regarding relationships with acquaintances' are formulated as core aspects of the dimension 'social integration'.

On the activity level, the assessment of the ‚physical functioning/occupational functioning' has a special relevance in the context of rehabilitation (Coster, Haley, Andres, Ludlow, Bond, & Ni, 2004). Here it is assessed to what extent patients are able to execute activities of everyday life without difficulty (e.g. 'lifting an object off the ground'). In empirical studies, the ‚physical functioning/occupational functioning' has been shown to be the most important predictor of a patient's participation and of treatment success (Reed, Lux, & Bufka, 2005). As the development of instruments for the direct assessment of participation is just at its beginnings (Heinemann, 2005), aspects of the ability to function are to be considered one of the most important indicators of treatment success. In the IRES patient questionnaire, the ‚physical functioning/occupational functioning' is subdivided into the aspects 'physical ability to function' and 'job-related ability to function' and comprehensively assesses functional constraints of the patients.

4. Questions concerning measurement of change

Intervention effects (changes) from the perspective of affected individuals may be directly assessed by means of explicit change ratings (e.g. "How has your condition changed?" worsened – remained the same – improved), by comparing status assessments from the beginning of treatment with status assessments following treatment (= indirect measurement of change), or by comparing status assessments following treatment with a remembered pre-status (= quasi-indirect measurement of change). Mittag et al. (2012) have systematized advantages and disadvantages of the different methods for measuring change and have derived recommendations for their application in rehabilitation research. Direct measurement of change is applicable economically as it merely requires one time of assessment, and it directly corresponds to everyday clinical life ("So, did it help you?"). The latter may temptingly lead to the assumption that respondents' statements

are easy to interpret in terms of content. However, the question of change is considerably more cognitively demanding than the question of current well-being (Mittag et al., 2012). Respondents not only have to be able to remember their state of health at a given point in the past but they also have to relate it to their present state. This method of measurement of change is therefore susceptible to a range of biases e.g. by memory distortions ("recall bias"), by current well-being ("present state bias") as well as by the respondent's interests and expectations.

Another problem is that this method of change measurement provides neither indicators regarding the initial state nor indicators regarding current well-being. This impedes the interpretation of experienced change. The information that nothing had changed, for instance, is to be interpreted differently concerning a bad initial state than concerning an advantageous initial state. Furthermore, studies show that the results of direct measurement of change represent rather a general factor independently of the respective contents while specific effects appear with indirect measures of different endpoints (Meyer-Moock et al., 2010).

Indirect measurement of change is considered to be the gold standard in clinical change measurement. Here, the intervention outcomes are determined by means of the difference between two status assessments that are separate in time (pre-post). Thus, distortions through memory effects or an excessive influence of the current well-being are eliminated to a large extent. Hence, indirect change measurement is often ascribed better measurement accuracy than direct measurement of change. Moreover, it reflects changes in different areas multidimensionally and in a differentiated manner (Meyer-Moock et al., 2010). Nevertheless, the considerably higher effort may be considered a disadvantage leading to a rather infrequent application of the indirect method of change measurement especially in routine surveys.

Indirect change measurement is associated with a range of methodical problems too (cf. Mittag et al., 2012). The measurements should represent the average (typical) situation at the given point in time. However, in practice, this prerequisite is not necessarily given. Situational conditions present at only one of the two measuring points (e.g. an acute illness, family incidents, extreme work loads) may lead to considerable "noise" in the measurement. Another problem concerning the validity of the measurement results from a possible shift of the respondents' evaluation standard at the second measurement ("response shift"). Such a change can be a consequence of changed internal standards (e.g. through comparison with other patients) or of a change in valences e.g. in heavy cases of illness or disability. By all means, this can be a desired consequence of the rehabilitation intervention, yet it may possibly lead to an over- or underestimation of the intervention's effects.

Just as the direct method, quasi-indirect measurement of change also requires only one measurement. Together with the assessment of the current state of health at the post-measure, well-being at the point of pre-measure is retrospectively assessed here and change is determined

by means of the difference between the two responses. Thus, the economy of direct measurement of change shall be joined with the advantages of indirect change measurement. Nonetheless, methodical problems remain here as well (e.g. „recall bias“). Direct and quasi-indirect methods of change measurement apparently tend to lead to an overestimation of effects compared to the direct method (Mittag et al., 2012). Simple recommendations for the implementation of a particular method of change measurement can not be provided. Depending on the question and the research design, the advantages and disadvantages of the different procedures have to be weighed.

5. Goal oriented measurement of change

The consideration of several parameters in the context of multidimensional and differentiated diagnostics results in a problem regarding the explanatory power of effect sizes for documenting treatment success: Generally, particular criteria are only relevant for some of the patients. No detectable changes in areas in which the patient was not strained at the time of admission or in which no changes were strived for by the treatment, can not be evaluated in terms of an insufficient effectiveness of rehabilitation. If, for instance, the total score of the IRES reflecting an aggregation of all IRES dimensions is used as the evaluation criterion, then changes important to the patient are possibly diluted or levelled out by an abundance of irrelevant information (Gerdes, 2005). Moreover, the more heterogeneous the original conditions and needs of the patients in the investigated sample, the smaller the detectable effects on the level of the individual dimensions. Heterogeneity on the level of variables and patients thus leads to a lack of validity concerning global change indexes and a systematic underestimation of relevant changes.

In order to solve this problem, Gerdes (1998) proposed the procedure of ‘Goal Oriented Measurement’ (German: Zielorientierte Ergebnismessung, ZOE) which, similar to the better known ‘goal-attainment-scaling` (GAS; Kiresuk, Smith, & Cardillo, 1994), enables individualized evaluation. Here, for every patient, only those scales of a multidimensional assessment instrument are considered in the evaluation that are at the beginning of treatment marked as a goal dimension in therapist-patient-dialog or in which the patient exhibits conspicuities. In calculating the ZOE total score, only those scales are used which have been defined as target areas (patient-related ZOE total score). On the level of the sample, for the evaluation of change on an IRES scale only those patients are considered who have marked the respective area as a target (scale- related ZOE total score): To what extent a rehabilitation intervention is suitable, for instance, to reduce problems in the realm of ‘stress’ is then only determined including those patients who wanted to focus on a reduction of this problem in the course of their treatment.

The ZOE-procedure has been applied in rehabilitation in several major evaluation studies (see Jäckel, Bengel, & Herdt, 2006). Despite its high apparent plausibility, there are also problems with the procedure (Zwingmann, 2002): The reliability and validity of the goal-setting, i.e. choosing the relevant problem areas, are unknown. Particularly, it is unclear to what extent the patient or the attending therapist or exactly their interaction is critical for choosing targets and whether the patient's problems or rather the potentially available range of treatments determines the choice. In addition, changes in the relevance of problem areas in the course of treatment are not considered (Beutler & Hamblin, 1986). Possible reactivity, in terms of a change in treatment organization due to the explicit determination of goals, is also not accounted for (Schulte-Bahrenberg, 1990). The ZOE-procedure is moreover blind toward unintended declines in non-targeted problem areas. One-sided focussing on the patient's goals may for instance lead to a neglect of therapy standards as they are defined e.g. in guidelines. This neglect, in turn, may lead to less good justification of the therapeutic action. Focussing on problem areas results in extreme group selection, so the validity of the treatment effectiveness evaluation is impacted by the phenomenon of regression to the mean (Zwingmann & Wirtz, 2005). As chance variability in extreme groups causes an overestimation of treatment-contingent changes, the interpretability of empirical effect sizes without suitable control groups is exacerbated. Further, as the ZOE total score for every patient is comprised of different individual aspects, the metric of the score is individually different and ultimately not interindividually comparable: The added indicators (individual areas marked as goals), for instance, each have a specific reliability and sensitivity to change.

Thus, the patient-specific ZOE total score is not only contentually incommensurable between patients but it also possesses variable statistical characteristics.

In the ZOE-procedure, constantly different partial samples whose composition is not controlled by the investigator are evaluated by means of different success measures. Further, the changes are systematically distorted by chance processes. This illustrates that using the ZOE-procedure the derivation of unequivocal statements regarding the effectiveness of treatment measures has to be taken with a grain of salt.

6. Item-Response-Theory and computer-supported adaptive testing

For the most part, in rehabilitation very comprehensive assessment-batteries are applied to meet the demand of evaluating state of health and aspects of health-related quality of life as comprehensively as possible. It must, however, be considered that data collection should be pinpointed against the background of the expected diagnostic benefit. One the one hand, it is not justifiable to have patients answer questions that are diagnostically irrelevant. And on the other hand, it has to be avoided that patients are strained too much by surveys resulting in subjective

responses that are unmotivated and have limited validity. To meet this demand, short forms or screening versions have been developed for assessment instruments. These are suitable to economically assess the respective crucial dimensions (e.g. SF-12, Ware et al., 1996; IRES-24, Wirtz et al., 2005; Kroenke, Spitzer, Williams, & Lowe, 2009). Due to few items of medium difficulty, however, these short forms often hold the problem that satisfactory diagnostic precision is reached merely in the medium range of strain. Assessing low or high strained patient groups with these instruments often results in unacceptably high standard errors as well as floor- and ceiling-effects in the sample distribution (McHorney, 1999).

An optimal combination of economical data collection and high measuring precision in all areas of the to be measured aspect continuum is yielded by computer-supported adaptive testing (CAT) based on Item-Response-Theory (van der Linden & Glas, 2000; Frey et al., 2011). Using CAT, the patient is requested to answer a question of medium difficulty at the beginning of the assessment of an aspect dimension (e.g. physical ability to function). Based on his/her answer, the next question possessing the maximum informative value is chosen from an underlying item-base. For example, if a patient answers the first question stating that he/she has problems covering a distance of 100 m by foot, then the question whether he/she can comfortably execute a day-long hiking trip is obsolete. It would then make more sense to inquire about the extent of problems regarding an easier activity, e.g. to what extent he/she has problems to independently move around their apartment. The algorithm underlying the CAT estimates the patient's ability after every answer and determines a degree of accuracy of this estimation (estimation error). The algorithm chooses questions to be answered until the estimation error for the patient's ability-level falls below a given limit.

Contrary to static test forms, similarly precise estimations are reached in all areas of ability as the achievable estimation accuracy is homogeneous in the whole spectrum of characteristics. Generally, responding to 6 to 8 items is sufficient for the ability estimation, so the strain on the patient in terms of time is reduced by more than 60% despite increased measuring quality. In the past years, CAT could already be successfully internationally implemented in many clinical projects (e.g. Haley, Fragala-Pinkham, Ni, Skrinar, & Corzo, 2005; Ware, Gandek, Sinclare, & Bjorner, 2005). Comprehensive information regarding CAT-supported assessment of the ICF-construct 'activities' may be retrieved on the internet at this address: ‚http://icfmeasure.com/'. A one-dimensional item-base has to be available as a prerequisite for CAT. An item-base consists of a group of items assessing the same aspect in differing degrees of difficulty, thus being strictly one-dimensional (Abberger, Haschke, Krense, Wirtz, Bengel, & Baumeister, in press; Hambleton, 1993; Haschke, Abberger, Müller, Wirtz, Bengel, & Baumeister, in press).

References

1. Abberger, B., Haschke, S., Krense, C., Wirtz, M., Bengel, J., & Baumeister, H. (in press). Development and calibration of an item bank for the assessment of anxiety in cardiovascular patients using Rasch analysis. Journal of Clinical Epidemiology.
2. Bernhard, J., Lowy, A., Mathys, N., Herrmann, R., & Hürny, C. (2004). Health related quality of life: A changing construct? Quality of Life Research, 13, 1187-1197.
3. Biefang, S., Birkner, B., Thien, U., Härtel, U., & Bullinger, M. (1997). Harmonisierung der Messung von Outcomes, Prädiktoren und Kosten sowie Prüfung geschlechtsspezifischer Unterschiede in der rehabilitationswissenschaftlichen Forschung. Rehabilitation, 36, 1-11.
4. Beutler, L.E. & Hamblin, D.L. (1986). Individual outcome measures of internal change: Methodological considerations. Journal of Consulting and Clinical Psychology, 54, 48-53.
5. Bruyere, S., Van Looy, S., & Peterson, D. B. (2005). The International Classification of Functioning, Disability and Health: Contemporary literature overview. Rehabilitation Psychology, 50, 113–121.
6. Bührlen, B., Gerdes, N., & Jäckel, W. H. (2005). Entwicklung und psychometrische Testung eines Patientenfragebogens für die medizinische Rehabilitation (IRES-3). Rehabilitation, 44, 63-74.
7. Bullinger, M. (2000). Erfassung der gesundheitsbezogenen Lebensqualität mit dem SF-36-Health Survey. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz, 43, 190-197.
8. Bullinger, M. & Ravens-Sieberer, U. (2000). Indikatoren des Rehabilitationsergebnisses. In J. Bengel & U. Koch (Hrsg.), Grundlagen der Rehabilitationswissenschaften (S. 305-322). Heidelberg: Springer.
9. Coster, W.J., Haley, S.M., Andres, P.L., Ludlow, L.H., Bond, T.L.Y., & Ni, P. (2004). Refining the conceptual basis for rehabilitation outcome measurement. Personal care and instrumental activities domain. Medical Care, 42, 62-72.
10. Deck, R., Mittag, O., Hüppe, A., Muche-Borowski, C., & Raspe, H. (2007). Index zur Messung von Einschränkungen der Teilhabe (IMET) – Erste Ergebnisse eines ICF-orientierten Assessmentinstruments. Praxis Klinische Verhaltensmedizin und Rehabilitation, 76, 113-120.
11. DIMDI (2004). ICF Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit. Köln: DIMDI.
12. Freund, A. M. & Ziegelmann, J. P. (2009). Lebensqualität: Die Bedeutung der Selektion, Optimierung und Kompensation. In J. Bengel & M. Jerusalem (Hrsg.), Handbuch der Gesundheitspsychologie und der Medizinischen Psychologie (S.475-48). Göttingen: Hogrefe.

13. Frey, C., Zwingmann, C., Böcker, M., Forkmann, T., Kröhne, U., Müller, E., & Wirtz. M. (2011). Adaptives Testen in der Rehabilitation - ein Weg zur ökonomischen Erhebung von Patientenmerkmalen. Rehabilitation, 50(3), 195–203.
14. Gerdes, N. (1998). Rehabilitationseffekte bei ‚Zielorientierter Ergebnismessung'. Ergebnisse der IRES-ZOE-Studie 1996/97. Deutsche Rentenversicherung, 3-4, 217-237.
15. Gerdes, N. (2005). Ein ICF-basiertes Theoriemodell der Rehabilitation als theoretischer Bezugsrahmen für den IRES-Fragebogen In L. Leonhart & N. Gerdes (Hrsg.), Der Einsatz des IRES-Fragebogens in der Rehabilitation (S. 93-110). Regensburg: Roderer.
16. Haley, S.M., Fragala-Pinkham, M.A., Ni, P.S., Skrinar, A.M., & Corzo, D. (2005). An adaptive testing approach for assessing physical functioning in children and adolescents. Developmental Medicine and Child Neurology, 47, 113-120.
17. Hambleton R. (1993). Principles and selected applications of Item Response Theory. In R. Linn (Ed.), Educational Measurement (pp. 147-200). Phoenix: Oryx.
18. Haschke, A., Abberger, B., Müller, E., Wirtz, M., Bengel, J., & Baumeister, H. (in press). Calibration of an item bank for work capacity in cardiological rehabilitation patients. European Journal of Preventive Cardiology.
19. Heinemann, A.W. (2005). Putting outcome measurement in context: A Rehabilitation Psychology perspective. Rehabilitation Psychology, 50, 6-14.
20. Hunt, S. M., McKenna, S. P., McEwen, J., Williams, J., & Papp, E. (1981). The Nottingham Health Profile: Subjective health status and medical consultations. Social Science & Medicine, 15A, 221-229.
21. Jäckel, W.H., Bengel, J., & Herdt, J. (Eds.) (2006). Research in rehabilitation. Results of a German Rehabilitation Research Network (Freiburg/Bad Säckingen). Stuttgart: Thieme.
22. Keith, R. A., Granger, C. V., Hamilton, B. B., & Sherwin, F. S. (1987). The Functional Independence Measure: A New Tool for Rehabilitation. In M. G. Eisenberg, R. C. Grzesiak (Eds.), Advances in Clinical Rehabilitation (pp. 6-18). Springer: New York.
23. Kiresuk, T.J., Smith, A., & Cardillo, J.E. (1994). Goal attainment scaling: applications, theory and measurement. Hillsdale: Erlbaum.
24. Kroenke, K., Spitzer, R.L., Williams, J.B.W., & Lowe, B. (2009). An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ-4. Psychosomatics, 50, 613-621.
25. McHorney, C. A. (1999). Health status assessment methods for adults: Past accomplishments and future challenges. Annual Review of Public Health, 20, 309-35.
26. Meyer-Moock S., Moock J., Mittag O., & Kohlmann T. (2012). The factor structure of direct and indirect methods for measuring change in medical rehabilitation--analyses on item level [in German]. Die Rehabilitation, 51, 118-128.

27. Mittag, O., Kohlmann, T., Meyer, T., Meyer-Moock, S., Meffert, C., Farin-Glattacker, E., Gerdes, N., Pohontsch, N., Moock, J., Jelitte, M., Löschmann, C., Bitzer, E.M., & Raspe, H. (2012). Empirically Derived Recommendations for Measuring Patient-Reported Change in Rehabilitation Studies [in German]. Die Rehabilitation (online-first; DOI: 10.1055/s-0032-1314876).
28. Peterson, D. B. (2005). International Classification of Functioning, Disability and Health: An introduction for Rehabilitation Psychologists. Rehabilitation Psychology, 50, 105–112.
29. Reed, G.M., Lux, J.B., & Bufka, L.F. (2005). Operationalizing the International Classification of Functioning, Disability and Health in Clinical Settings. Rehabilitation Psychology, 50, 122-131.
30. Schulte-Bahrenberg, T. (1990). Therapieziele, Therapieprozess und Therapieerfolg. Pfaffenweiler: Centaurus.
31. Schulz, H. (2002). FIM Manual. Messung der Funktionalen Selbständigkeit (Functional Independence Measure). Meerbusch: Schmidt.
32. Szcabo, S., on behalf of the WHOQOL Group (1996). The World Health Organization Quality of Life (WHOQOL) Assessment Instrument. In: B. Spilker (Ed.), Quality of life and pharmacoeconomics in clinical trials (pp. 355-362). 2nd ed. New York: Lippincott-Raven Publishers.
33. Van der Linden, W.J., & Glas, C.A.W. (2000). Computerized adaptive testing: Theory and practice. Boston: Kluwer.
34. Ware, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey. Medical Care, 34(3), 220-233.
35. Ware, J. E., Gandek, B., Sinclair, S. J. & Bjorner, J. B. (2005). Item response theory and computerized adaptive testing: Implications for outcomes measurement in rehabilitation. Rehabilitation Psychology, 50, 71-78.
36. Ware, J. E. & Sherbourne, C. D. (1992). The MOS 36-Item Short-Form Health Survey (SF-36). Conceptual framework and item selection. Medical Care, 30(6), 473-483.
37. Wirtz, M., Farin, E., Bengel, J., Jäckel, W.H., Hämmerer, D., & Gerdes, N. (2005). IRES-24 Patientenfragebogen – Entwicklung der Kurzform eines Assessmentinstrumentes in der Rehabilitation mittels des Mixed-Rasch-Modells. Diagnostica, 51, 75-87.
38. Zwingmann, C. (2002). Der IRES-Patientenfragebogen. Psychometrische Reanalysen an einem rehabilitationsspezifischen Assessmentinstrument. Regensburg: Roderer.
39. Zwingmann, C. & Wirtz, M. (2005). Regression zur Mitte. Rehabilitation, 44.

Анализ эффективности реабилитации: результаты, оценка и измерение изменений

Дж. Бенгель,
Университет Альберта-Людвига, Фрайбург (Германия)

М. Крафт,
Университет Альберта-Людвига, Фрайбург (Германия)

О. Миттаг,
Университетский медицинский центр г. Фрайбурга (Германия)

Оценка эффективности реабилитационных вмешательств крайне важна в реабилитационной практике и ее исследованиях. Данная статья посвящена наиболее важным аспектам измерения успеха реабилитационной работы и связанных с ней изменений. В ней приводятся как ключевые критерии и индикаторы результативности, так и наиболее распространенные инструменты оценки, основанные на самоотчетах клиентов. Один из них – опросник IRES-3 (Индикаторы реабилитационного статуса), разработанный в Германии – рассматривается более подробно. Обсуждаются вопросы измерения изменений и его альтернативные методики, такие как целеориентированные измерения (goal-oriented measure) и адаптированное тестирование, основанное на стохастической теории тестов (Item-Response-Theory).

Ключевые слова: диагностические измерения, методы оценки, оценка результатов, качество жизни, стохастическая теория тестов, реабилитация.