

# Patient-Report-Outcome-Measure and Incentives for Inpatient Chronic Care in Germany

Tobias Romeyke<sup>1,2</sup>, Elisabeth Noehammer<sup>1</sup> & Harald Stummer<sup>1</sup>

<sup>1</sup> UMIT – University for Health Sciences, Medical Informatics and Technology, Institute for Management and Economics in Health Care, Austria

<sup>2</sup> Waldhausklinik, Acute Hospital for Internal Medicine, Pain Therapy, Complementary and Individualized Patient Centred Medicine, Deuringen, Germany

Correspondence: PD Dr. Tobias Romeyke, University for Health Sciences, Medical Informatics and Technology, Institute for Management and Economics in Health Care, Eduard Wallnöfer-Zentrum 1, 6060 Hall in Tirol, Austria. Tel: 43-50-8648-3871. E-mail: tobias.romeyke@ext.umat.at

Received: March 25, 2020 Accepted: May 28, 2020 Online Published: June 12, 2020

doi:10.5539/gjhs.v12n8p127

URL: <https://doi.org/10.5539/gjhs.v12n8p127>

## Abstract

**Introduction:** In general, incentive tools like pay for performance (P4P) have positive effects regarding treatment quality and financial outcomes. As they are applicable to the clinical management of chronic conditions like asthma and diabetes, this article analyses their potential for multimodal complex treatment of chronic rheumatic diseases.

**Methods:** Cost data for chronic rheumatic diseases with and without specified complex treatments and their respective remuneration are compared to permit specific statements regarding incentive aspects in a DRG setting. Moreover, several standardized Patient-Report-Outcome-Measures (PROMs) are considered in the context of complex treatment to allow not only for efficiency, but also effectiveness evaluation.

**Results:** In total, 375 patients with complex treatment for rheumatic conditions were surveyed from 2013 to 2018. The incentive is slightly below (4,821.05 €) the costs incurred for complex treatments (4,972.44 €). The results of the used PROMs are positive as pain intensity decreased considerably ( $p < .001$ ,  $r=0.75$ ) and mental state complaints were reduced ( $p < .001$ ).

**Conclusions:** PROMs are valid instruments to capture changes in patient well-being. They also help to improve clinical operations and can be used for benchmarking. The P4P approach should cover the costs incurred to ensure the incentive structure.

**Keywords:** chronic care, incentive, multimodal treatment, patient-report-outcome-measure, pay for performance, rheumatic and musculoskeletal diseases

## 1. Introduction

It has been estimated that around 2 billion people worldwide suffer from rheumatic diseases (Vos et al., 2012). More than 200 of these conditions are currently known (Van Der Heijde et al., 2018). Many of these are chronic and affect patients' quality of life (Anyfanti et al., 2016; Van Der Heijde et al., 2018) and the cost structure of the healthcare system for a long time.

Healthcare policy makers and service providers are therefore required to provide for adequate and efficient care. While efforts to improve the quality of care have long been debated, service providers are offered few incentives to improve the quality delivered due to limited resources. This reduces the likelihood of implementation of innovative and improved care management. Being problem-oriented and applied to the specific patient (DiPiero & Sanders, 2005), it is more effective, but more costly.

Incentive tools like pay-for-performance (P4P) or pay-for quality (P4Q) are discussed and analysed worldwide in this context (Van Herck et al., 2010; Ryan & Doran, 2012). The Quality and Outcomes Framework (QOF) currently introduced in the UK is one of the key P4P projects (Ryan & Doran, 2012). P4P is linked to the quality of structure, processes, and results. Thus, health service is provided in accordance with specified quality criteria while taking into account the financial burden (Ryan & Blustein, 2012). In general, it can be assumed that

optimized structure and processes will have an impact on the quality (Donabedian, 1966).

However, the implementation of P4P programs is often difficult because incentive structures designed to achieve quality targets and improve care structures are complex (Kondo et al., 2016; Mendelson et al., 2017). This is one of the reasons why P4P are often disease-related and only focus on specific diseases. Thus, they cannot adequately take into account individual needs of chronically ill, multimorbid patients (Boyd et al., 2005; Barnett et al., 2012). Inflexible linking of P4P programs and clinical practice guidelines or quality initiatives has been the subject of critical discussions (Commonwealth Fund international health policy survey, 2010).

As instruments for specific indications, P4P initiatives are being used in the treatment of chronic diseases such as asthma and diabetes (Levin-Scherz et al., 2006; Scott et al., 2009). Extensive analyses in both outpatient and inpatient settings tend to show positive effects of P4P in the treatment of these (Van Herck et al., 2010), suggesting the utilization of P4P for further chronic conditions like rheumatic diseases. Nevertheless and despite the trend to develop P4P instruments for specific indications, remuneration schemes that are linked to quality have not been widely used in the field of rheumatology (Harrison et al., 2016). In addition, their effectiveness in promoting higher quality of care needs scientific inquiry and evaluation (Bombardier & Mian, 2013). The aim of this article is thus twofold. Taking the German DRG system as basis, which allows distinguishing between treatments of rheumatic diseases with and without multimodal complex treatment (MRT), we analyse the P4P effect of MRT and its effectiveness based on Patient Reported Outcomes. Moreover, we add to the literature on P4P in the context of multimorbidity. MRT cover a very broad range of individualized interventions and their combination because of the diverseness of the illness. This mimics the situation of chronically ill, multimorbid patients.

### *1.1 The Multimodal Rheumatologic Complex Treatment (OPS 8-983) (MRT)*

Hospitals receive different remunerations depending on whether they treat (rheumatic) diseases in a classical way or using complex treatments. For the latter, specific treatment codes (so called complex treatment codes) were developed to make costly and specialised treatments billable in the German diagnosis related groups (G-DRG) system. With regard to the remuneration and the costs incurred, cost-unit accounting determines the total costs of complex treatment for each patient. In DRG contexts, payments for cases are flat-rates calculated based on the average base rate times a cost weight.

This study investigates the multimodal rheumatologic complex treatment, which is classified in the German procedure classification (OPS) (OPS 8-983). This complex treatment was selected because it ensures holistic care of patients with chronic rheumatic conditions according to guidelines for quality assurance of inpatient services. The German Institute of Medical Documentation and Information (DIMDI) publishes the OPS procedure classification on behalf of the German Federal Ministry of Health. It lists all complex treatment codes. OPS is an adaptation of the International Classification of Procedures in Medicine (ICPM) of the World Health Organisation (WHO).

An implementation of MRT places stringent requirements on the structure and procedural organization of a hospital and requires specialist rheumatologic supervision. Specific quality criteria that concern the structure, process and result quality must be ensured for the MRT. The specification of head physician qualifications, provision and implementation of physiotherapy, occupational therapy, pain treatment, cognitive behavioural therapy, and psychotherapy are structural quality criteria. At least three of the following treatment modes must be used: “physiotherapy”, “occupational therapy”, “pain treatment”, “cognitive behavioural therapy”, and “person-centred psychotherapy”. Achieving a specific treatment density per week and performing assessments at a specific time relate to process quality. At least 11 hours of treatment per week are required per patient. An immediate start of multimodal treatment also affects processes. It must be guaranteed that pain treatment and physiotherapy are started without delay. Outcome quality indicators are defined by specifying assessment parameters. In addition to a standardized assessment, it is also necessary to measure the disease activity, functional impairment, and the pain extent at the beginning and end of hospitalization. Treatment results are evaluated according to patients’ assessment of changes in pain intensity, mental state, subjective evaluation of impairment due to physical and general complaints, and physical mobility. The examination of outcome quality should provide information about the effect of structure and process quality on treatment results. Furthermore, these plus the costs incurred need to be compared with the higher remuneration for the performed service.

### *1.2 Patient-Report-Outcome-Measure (PROM)*

Despite advances in the pharmacological treatment of rheumatic diseases and its positive impact on disease activity and inflammatory parameters, patients often report that their condition negatively affects their well-being due to symptoms such as pain, fatigue, morning stiffness, or depression (Fautrel et al., 2018). Medical technology

cannot or not sufficiently measure many patient-related data. This concerns, for example, depression, perceived strength or frequency of symptoms (such as headaches), or the effect of the disease on performing everyday tasks.

Data from routine self-assessment of a patient's state of health, called Patient-Report-Outcome-Measure (PROM), alleviate this issue. PROMs allow patients to answer questions about their state of health independently, without any interference from the outside. Patients record the impact of their illness from their own point of view (Dawson et al., 2010), which together with medical data leads to a comprehensive assessment of treatment benefits (Black 2013). This is helpful to better focus on the needs of the individual patient (Greenhalgh et al., 2017), to treat the symptoms more efficiently (Food and Drug Administration, 2009; Chin & Lee, 2008), and to monitor the course of treatment (Rotenstein et al., 2017). Furthermore, Patient Reported Outcomes (PROs) have the potential to improve patient-physician communication and to increase satisfaction through active patient involvement (Valderas et al., 2008). From a scientific point of view, PROs can help to better account for patient perspectives when taking treatment decisions (Haywood, 2006). This creates the basis for patient-oriented care.

However, even though they can help to make precise, efficient and flexible (Cella et al., 2010) statements about the quality of treatment (Castrejon & Pincus, 2012), quality measurement instruments such as PROM are rarely used for patients with chronic conditions. By including standardized PROMs, this article helps closing this research gap. The next section describes the approach employed.

## 2. Material Studied

In total, 375 patients were surveyed using PROMs to assess their state of health on admission and on discharge (see 3.1.). The retrospective study includes all patients with close to complete datasets on PROMs. Missing data thus does not represent a sufficient risk of bias (Sloan et al., 2007). The primary diagnosis was confirmed by a specialist in rheumatology and identified, according to the German coding guidelines, as "the diagnosis which, after analysis, has been established as the principal reason for admitting the patient" (Deutsche Kodierrichtlinien). Only patients with degree 2 and 3 of chronicity were included. The basis was the Mainz Pain Staging System (MPSS), which considers temporal aspects (pain progression), spatial aspects (pain localization), medication-taking habits, and occupation of the patient. The MPSS is considered a valid measuring instrument (Frettlöh et al., 2003, Wurmthaler et al., 1996). This created a homogeneous group of patients that could be used to evaluate the PROMs and to calculate treatment costs.

All patient data is anonymised as only the diagnosis, degree of chronicity and PROM outcomes are of relevance for the analysis. The study was approved by the Research Committee for Scientific Ethical Questions of UMIT - University of Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria.

## 3. Methods

Descriptive and inductive statistical calculations were generated using the statistical software package SPSS Statistics 25.0 (IBM, 2017). We employed dependent t-tests in cases of normal distribution and the Wilcoxon test to verify if there are significant differences between two measurements where normal distribution was violated. This method is also very resistant to outliers and uses two paired samples to check whether the central tendencies of the underlying paired populations are equal (Siegel & Castellan, 1956). Below, the PROMs and cost data analysed are described.

### 3.1 Patient-Report-Outcome-Measure (PROM)

Patients were surveyed using PROMs on admission and on discharge from the hospital. The MRT specifies requirements for staff qualification, evidence-based use of treatment modes, treatment intensity, processes, and surveying by means of PROs. It does not yet prescribe which PRO values should be achieved until the patient is discharged. This section presents the instruments used in this study.

#### 3.1.1 Visual Analogue Scale (VAS)

This study employed the VAS to measure pain. VAS has been widely used for many years, also for patients with rheumatic conditions (Downie et al., 1978). It can be done in less than a minute (Downie et al., 1978), is therefore economical and requires few resources plus is well accepted by patients (Huskisson, 1974). It consists of a 100 mm horizontal line on which respondents mark the point that matches the amount of pain they feel. "No pain" equals 0 and "unbearable pain" equals 10 on the scale (Burckhardt & Jones, 2003). The advantage of VAS is its very high degree of resolution (possibility of finest gradations) (Flynn et al., 2004). VAS values are rationally scaled and allow the application of parametric tests, assuming normal distribution. Reliability and validity of measurements with VAS have been confirmed by many studies (Bond & Lader, 1974; Gift, 1989). The test-retest reliability is very high with  $r = .98$ .

### 3.1.2 Patient Health Questionnaire (PHQ)

The Patient Health Questionnaire (employed in the German Version PHQ-D) is used to evaluate patients with acute pain exacerbations due to a chronic rheumatic disease, as these patients often develop a depression. The PHQ-D, also known as “PHQ-9”, has established itself as a psycho-diagnostic tool in the clinical practice and science (Spitzer et al., 1992; Gilbody et al., 2007; Kroenke et al., 2001) not only in treatment, but also for follow-up (US Preventive Services Task Force, 2002). The tool can be used in inpatient settings (Stafford et al., 2007) and is recommended by the American Psychiatric Association working group. The following values can be assigned to answers: 0 (“not at all”), 1 (“on some days”), 2 (“on more than half of the days”), and 3 (“nearly every day”). Values below 5 indicate absence of a depressive disorder. Values between 5 and 9 correspond to a mild or persistent depressive disorder. Moderate depressive episodes are present at 10-14 points. 15 to 19 points indicate a moderately severe, and 20-27 points a major depressive disorder. A meta-analysis demonstrated high sensitivity and specificity of the studies, with 80% and 92%, respectively (Gilbody et al., 2007). The internal consistency of this tool is very high (Cronbach’s  $\alpha = 0.88$  and  $\alpha = 0.89$ ) (Gräfe et al., 2004; Kroenke et al., 2001). The same is true for its test-retest reliability.

### 3.1.3 Zerssen

The “Von Zerssen complaint list” is used to assess the extent of the subjective impairment caused by physical and general symptoms (Zerssen & Koelle, 1976; Zerssen & Petermann, 2011). It has been selected because it can be used in patients with advanced chronicity and comorbid mental conditions. The tool consists of 24 items that designate the respective symptoms. The scores of the individual items (0 = not at all, 1 = rarely, 2 = moderately, 3 = strongly) are added up. The internal consistency is  $\alpha = .94$ . The split-half reliability is high with  $r = .93$ . With regard to the validity, a value of  $r = .62$  results in correlation with the criterion of belonging to a clinical or control group (Fahrenberg et al., 2001).

## 3.2 Calculation of Treatment Costs

There are direct and indirect cost centres. Direct cost centres provide services that are directly related to the patient. These are nursing, examination, and treatment services. Indirect cost centres, on the other hand, provide services to direct cost centres without a relation to patients. These are, for example, “pharmacy”, “bed maintenance”, “provision of goods and services”, and “administration”. Costs of these indirect cost centres must be charged to the direct cost centres using appropriate distribution keys.

Costs are calculated in accordance with the specifications of the Calculation Manual (version 3.0) by the German Institute for the Hospital Remuneration System (InEK). This manual must be used for the further development of the German DRG system. Patient-related treatment cost calculations include all treatment cases, services, and costs of the hospital that fall under the DRG system’s compensation framework under the applicable regulatory requirements. This includes the collection of case-related data and cost data (approximately 41,250 data for this study). The calculation of the cost of treatment, for all cases, uses a total cost approach on an actual cost basis. In both – cost element accounting and cost centre accounting – the total hospital costs are adjusted for cost components that are not related to services included in the DRG system. In addition, a breakdown of costs should ensure consistency between cost and service volumes reported for each cost centre. Table 1 shows the key cost calculation steps of the total calculation. The key steps of cost centre accounting and cost unit accounting are summarized in Tables 2 and 3. Cost elements are grouped into cost element groups according to the specifications of the Calculation Manual (Table 4).

Table 1. Overview of the central calculation steps

Preparation of the data required for the calculation (case-related data, cost data, internal performance calculation billing key)
Personnel cost accounting
Determination of the relevant services
Calculation of Costs Relevant to Costs: Work within Cost Element Accounting
Calculation of Cost-Relevant Costs: Work in Cost Center Accounting
Cost center accounting
Carrying out cost object accounting (direct costs and overheads)

Table 2. Central steps of cost center accounting

The cost of indirect medical cost centers Infrastructure must be assigned to cost element group 7 of direct cost centers
Assignment of indirect cost centers not medical infrastructure to the cost element group 8 of the direct cost centers
Relief of the collection cost centers
Creation of a sum and balance list for the cost centers before the implementation of the internal performance calculation
Assignment of allocation keys for each indirect cost center of medical and non-medical infrastructure
Implementation of the internal performance calculation based on the chosen procedure
Creation of a sum and balance list after execution of the internal performance calculation
Creation of a sum and balance list for the accruals cost centers in each case before and after their discharge on the accrued expenses
Consolidation of cost elements per direct cost center to cost element groups

Table 3. Central steps of cost object accounting

Assignment of direct cost centers to cost center groups
Determination of the costs that are assigned as direct costs
Selection of suitable reference variables for case-related cost assignment for each direct cost center
Creation of calculation rates for direct cost centers
Distribution of overheads of direct cost centers to the benefit receiving cases using the calculation rates
Assignment of the direct costs according to the consumption documentation in the case data record, on the corresponding cost modules of the treatment cases
Compaction of cost data to cost modules

Table 4. Overview of the cost element groups of a treatment case

Cost Element Group (CEG) 1	Personnel costs, medical service
CEG 2	Personnel costs, nursing service
CEG 3	Personnel costs of the functional service and the medical-technical service
CEG 4a	Material costs for pharmaceuticals
CEG 4b	Material costs for pharmaceuticals (direct costs / actual consumption)
CEG 5	Material costs for implants and transplants
CEG 6a	Material costs of medical supplies (excluding pharmaceuticals, implants and transplants)
CEG 6b	Material costs of medical supplies (direct costs / actual consumption, excluding pharmaceuticals, implants and transplants)
CEG 7	Material costs of medical personnel and material costs of the medical infrastructure
CEG 8	Personnel and material costs of non-medical infrastructure

## 4. Results

### 4.1 Cost Analysis

The cost weight for remuneration of hospitals (reference range: years 2013-2018) for conventional treatment of fibromyalgia is, on average, 0.8945. If an average base rate of € 3245.63 is applied over this period, the hospital

will receive a flat-rate sum of € 2903.22. Treatment of rheumatoid arthritis is remunerated with € 2661.42, with a cost weight of 0.820. Treatments of ankylosing spondylitis and polymyalgia rheumatica are remunerated with € 2421.24 (Figure 1). If patients are treated within the scope of the MRT, taking into account the specified structural characteristics, cost bearers will pay € 4821.05 to the hospitals. This is independent of the specific rheumatic disease. The (potential) incentive portion is therefore on average € 2159.09 higher than for conventional treatments.

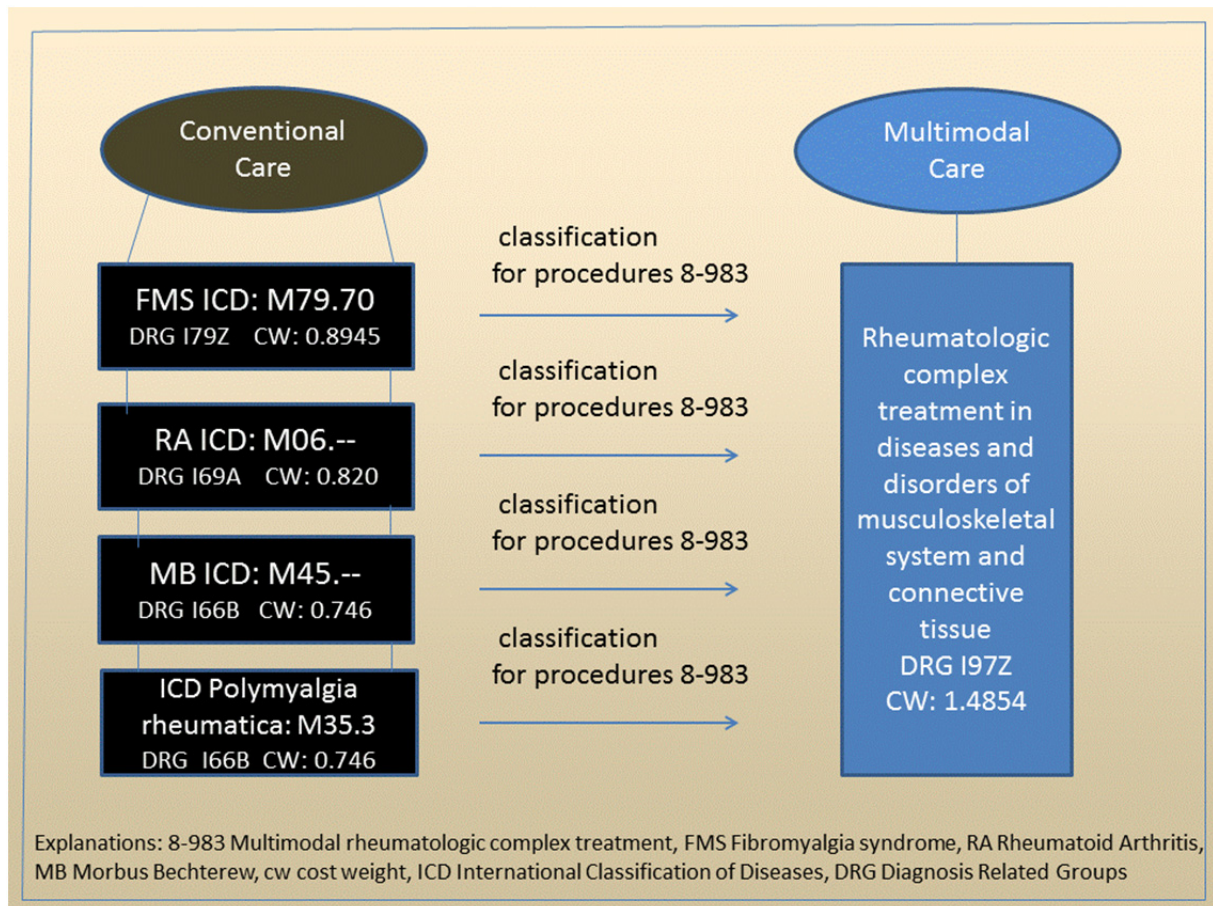


Figure 1. Value-based Purchasing: Conventional Care and Multimodal Care

#### 4.1.1 Cost-Unit Accounting

To calculate the incentive portion, a standardized cost-unit accounting was performed for the 375 treated patients including all cost types listed in Table 4 except for CEG 4b, 5, and 6b since there were no costs involved, or the amounts were very small. CEG 1, 2, and 8 account for the largest share of costs. CEG 1 includes salaries of doctors, social security contributions, and fees for non-salaried doctors. CEG 2 includes salaries and social security contributions of nursing staff. CEG 3 covers salaries and social security contributions of medical-technical staff. CEG 4a covers costs of medicaments. 6a covers dressings, therapeutic remedies, and aids. CEG 7 covers essentially the cost of patient transport and medical equipment maintenance. CEG 8 covers salaries of administration and technical staff, their social security contributions, and cost of electricity, food, and hospital admin. An average total cost of € 4,972.44 was calculated per patient for the multimodal rheumatic complex treatment. The incentive is thus slightly below the costs incurred (Figure 2).

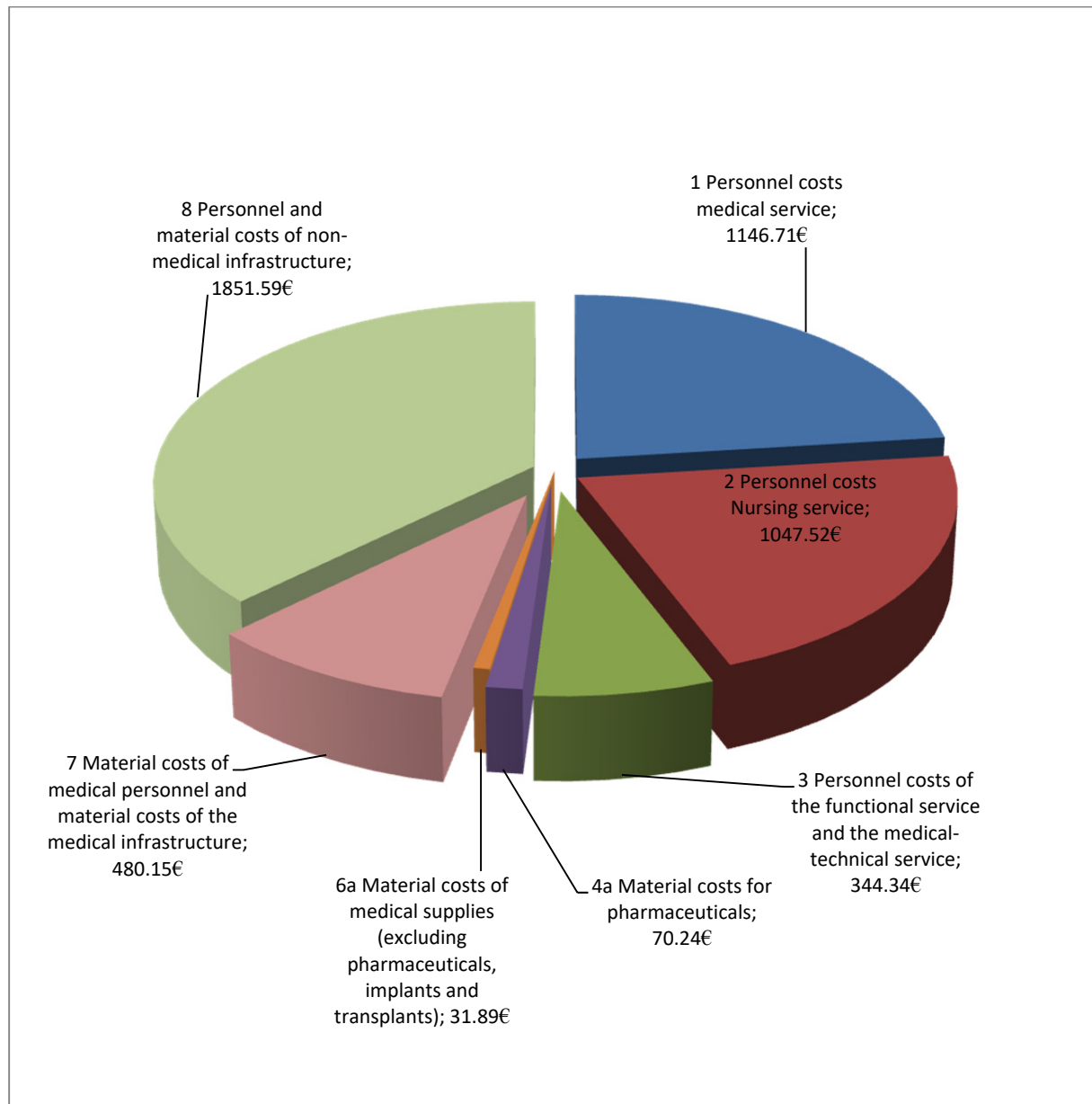


Figure 2. Results from the cost-unit accounting: multimodal rheumatic complex treatment

#### 4.2 Quality criteria and PROs

This study investigated 375 complex treatments. Patient data for pain intensity (is complete, no value is lacking) was collected by means of VAS at admission and discharge. The average pain intensity was reduced from 7.028 to 4.205 until the patient was discharged, as shown in Figure 3a. Figures 3 b and c show the evaluation of well-being at admission and discharge. A reduction of the value means the state improved, which is true for the majority of patients.

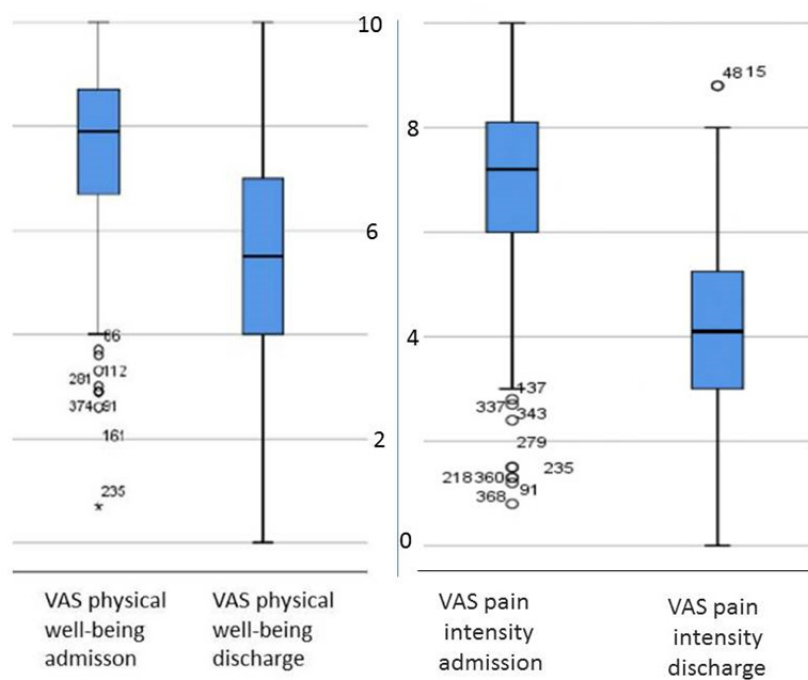


Figure 3a. Boxplots; VAS physical well-being and pain intensity on admission and discharge

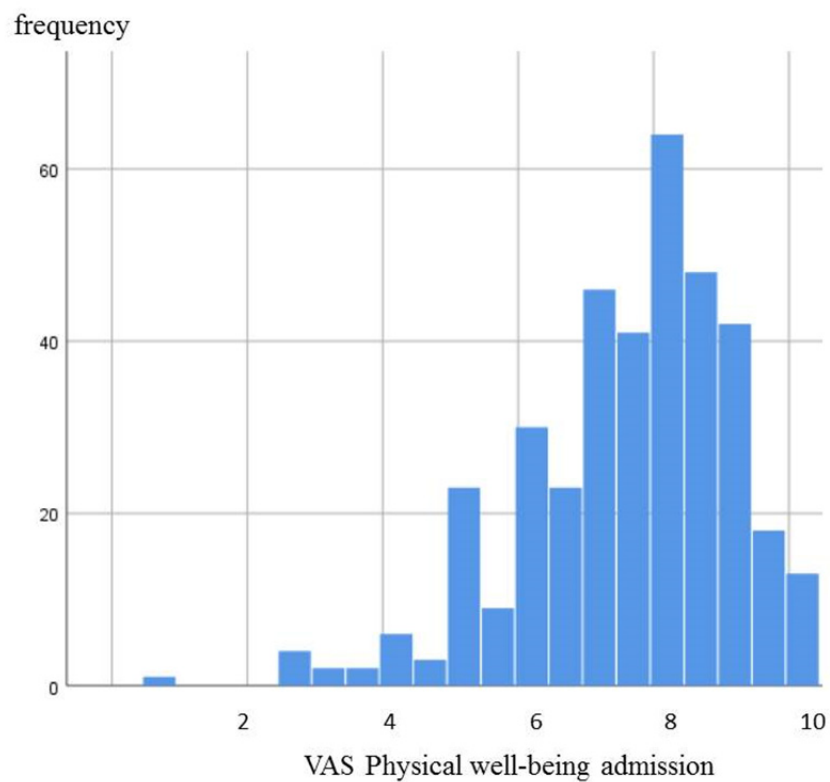


Figure 3b. VAS Physical well-being on admission



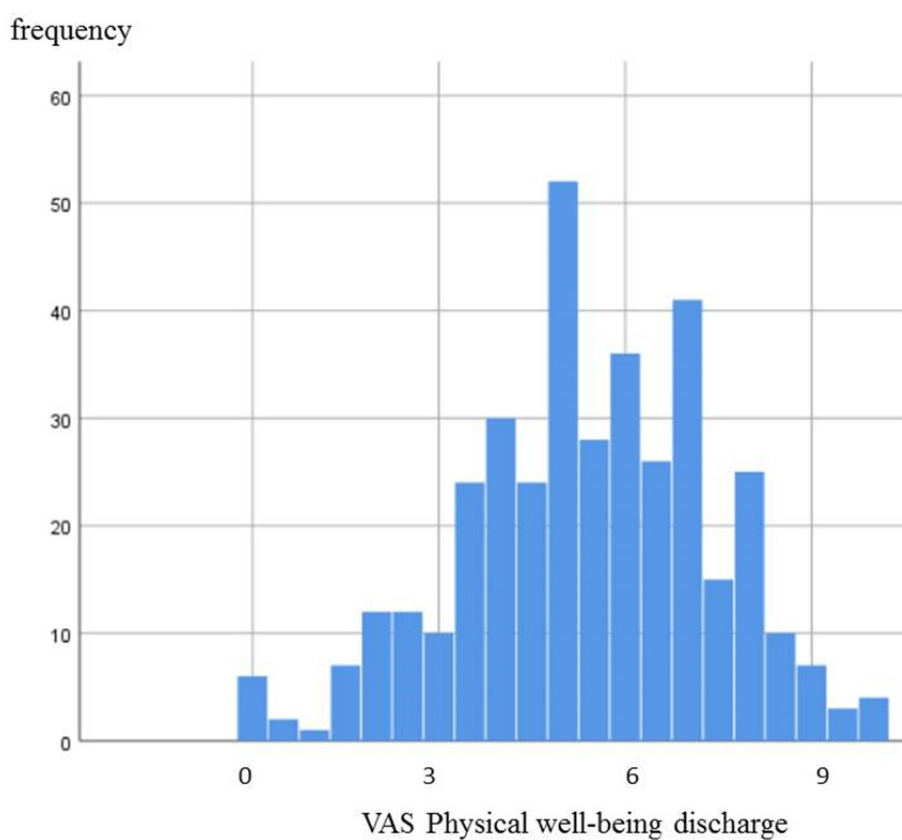


Figure 3c. VAS Physical well-being on discharge

The p-value is highly significant regarding changes in perceived pain (t-test,  $p < .001$ ) over time (Tables 5a, b).

Table 5a. Descriptive values (VAS) for both measurements

	Average	N	standard deviation	Standard error of the mean value
pain intensity on arrival (VAS)	7.028	375	1.7212	.0889
intensity of pain at discharge (VAS)	4.205	375	1.7074	.0882

Table 5b. Pain intensity: t-test for connected samples

	Paired differences					T	df	Sig. 2-sided
	Average	standard deviation	Standard error of the mean value	95% confidence interval of the difference				
				lower	upper			
pain intensity on arrival (VAS)	2.8232	2.2012	.1137	2.5997	3.0467	24.837	374	.000
intensity of pain at discharge (VAS)								

The values for VAS “Physical well-being” (Table 5c), also significantly differ between admission and discharge

(asymptotic Wilcoxon test:  $z = -14.514$ ,  $p < .001$ ,  $n = 375$ ). The effect size is  $r = 0.75$  and corresponds to a very strong effect.

Table 5c. VAS physical well-being

		N	Middle rank	Rank sum
	Negative ranks	317 <sup>a</sup>	196.75	62370.00
VAS physical well/being discharge	Positive ranks	47 <sup>b</sup>	86.38	4060.00
VAS physical well/being admission	Bonds	11 <sup>c</sup>		
	total	375		

a. VAS physical well/being discharge < VAS physical well/being admission.

b. VAS physical well/being discharge > VAS physical well/being admission.

c. VAS physical well/being discharge = VAS physical well/being admission.

Changes of mental state measured by means of PHQ-9 show a reduction of the value in the majority of patients, which means that their mental state has improved. This can be seen in the clear shift of values to the left in Figure 4b in comparison with Figure 4a, which means that the score value decreases.

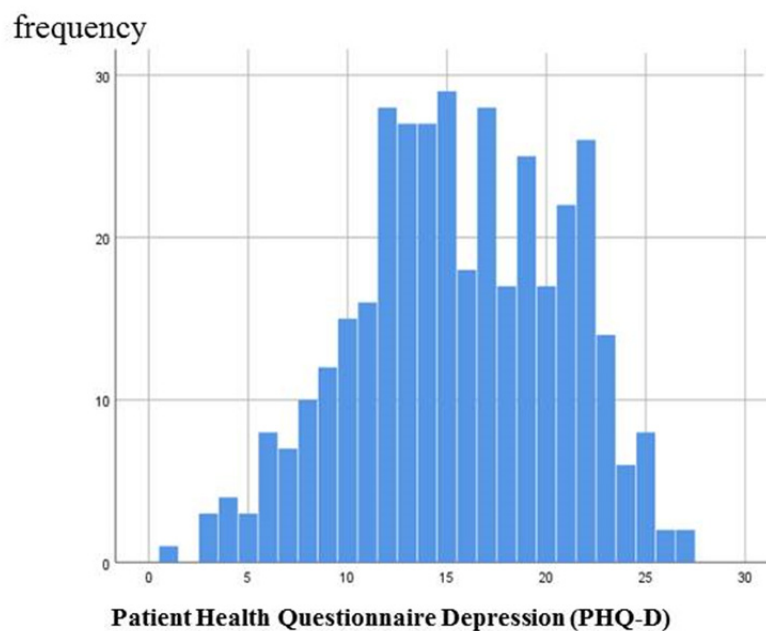


Figure 4a. Frequency of Patient Health Questionnaire Depression (PHQ-D) on admission

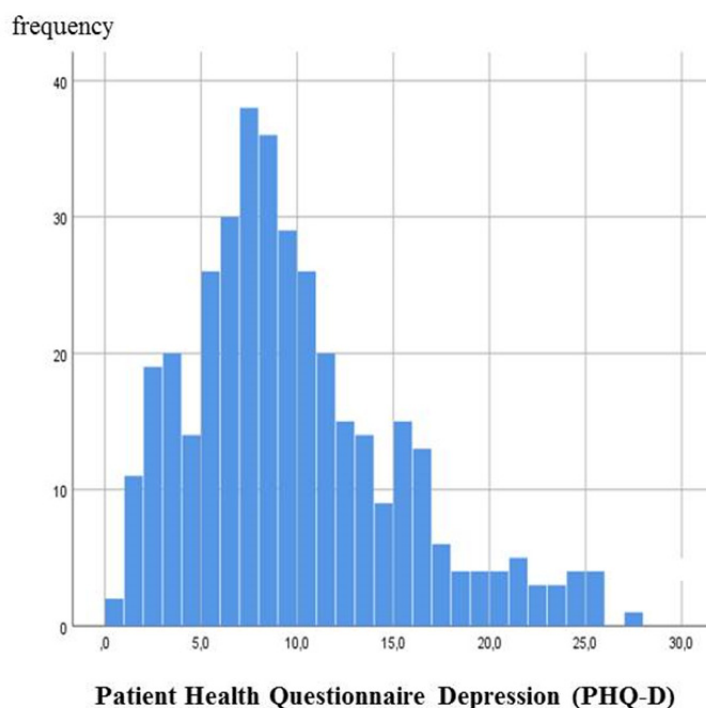


Figure 4b. Frequency of Patient Health Questionnaire Depression (PHQ-D) discharge

Overall, and based on a complete response rate ( $N = 375$ ), the values decrease highly significantly (WSR  $p < .001$ ) (Tables 6a, b).

Table 6a. PHQ-D discharge - PHQ-D on arrival

		N	Middle rank	Rank sum
PHQ-D discharge - PHQ-D on arrival	Negative ranks	328 <sup>a</sup>	189.05	62007.00
	Positive ranks	30 <sup>b</sup>	75.13	2254.00
	Bonds	17 <sup>c</sup>		
	total	375		

a. PHQ-D discharge < PHQ-D arrival.

b. PHQ-D discharge > PHQ-D arrival.

c. PHQ-D discharge = PHQ-D arrival

Table 6b. Wilcoxon signed-rank test PHQ-D discharge - PHQ-D on arrival

Test<sup>a</sup>

	PHQ-D discharge - PHQ-D arrival
Z	-15.258 <sup>b</sup>
Asymptotic significance (2-sided)	.000

a. Wilcoxon signed-rank test (WSR)

b. Based on positive ranks.

One patient assessment at discharge is missing from the assessment acc. to Zerssen. Overall, the negative ranks clearly outweigh the positive ones. The values decrease significantly, which indicates a reduction in complaints (WSR  $p < .001$ ) (Tables 7a,b).

Table 7a. Zerssen discharge - Zerssen on arrival

		N	Middle rank	Rank sum
Zerssen discharge - Zerssen arrival	Negative ranks	275 <sup>a</sup>	194.50	53488.00
	Positive ranks	75 <sup>b</sup>	105.83	7937.00
	Bonds	24 <sup>c</sup>		
	total	374		

a. Zerssen discharge < Zerssen arrival.

b. Zerssen discharge > Zerssen arrival.

c. Zerssen discharge = Zerssen arrival.

Table 7b. Wilcoxon signed-rank test Zerssen discharge - Zerssen on arrival

Test<sup>a</sup>

	Zerssen discharge - Zerssen arrival
Z	-12.028 <sup>b</sup>
Asymptotic significance (2-sided)	.000

a. Wilcoxon signed-rank test (WSR).

b. Based on positive ranks.

## 5. Discussion

For several years, there has been a discussion around performance-based pay for healthcare providers, such as hospitals and doctors, which should provide incentives for better quality and efficiency of care (Jha et al., 2012). This approach is also supported by the WHO (Richard, 2010). Complex treatments are an innovative construct that can be used to set binding structural, procedural, and outcome indicators for hospitals. However, their implementation is complex in itself. Regarding P4P, contradictory results are debated (Roberts et al., 2018). In the case of specific chronic diseases, there is also a risk that P4P could exclude patients from a treatment program (Simpson et al., 2007) should the incentive be too small to cover costs. Therefore, linking incentives for providers to qualitative structural characteristics and ensuring cost-coverage seems advisable for the treatment of multimorbid patients with chronic conditions.

So far, only a few studies have evaluated the outcome of complex treatment (Romeyke & Stummer, 2014; Klemm et al., 2019). In addition, analyses of incentive structure focusing on multimodal treatments for acute inpatient treatment of rheumatic diseases in conjunction with outcome parameters have not been investigated scientifically. The aim of this article was thus to analyse the costs, remunerations and PROM results of treatments for rheumatic diseases with and without multimodal complex treatment in the German DRG context. This was done to check for effectiveness (PROM) and potential incentives (higher remunerations) of multimodal complex treatments in rheumatic diseases, closing a research gap there.

Cost analyses demonstrate that the share of incentives is approximately 85% higher compared to a conventional treatment without statutory requirements. However, the results of cost calculation show (1) that this does not entirely cover the costs incurred and (2) a high cost share in CEG 8. In addition to salaries and social security contributions of administrative staff, providers of goods and services, and technical service staff, this cost type includes the entire food and electricity supply, and additional organizational effort required to coordinate the treatment team. Procedural requirements concerning MRT are very high as the coordination of the treatment team, the control of treatment intensity, integrating PROMs, and the control of length of stay have to be covered. An improved process management of complex treatments might help to close the remuneration gap. Specially trained assistants could provide the treatment team with support regarding planning of schedules, etc. On a systems and

societal level, interventions regarding prevention of chronic conditions could be improved to reduce the number of patients requiring complex treatment. The increasing demand for individual and holistic treatment programs for patients with long-term rheumatic and musculoskeletal diseases opens a new chapter in rheumatology (Nikiphorou et al., 2018). With increasing competition for patients, hospitals will seek to improve patient satisfaction. In the future, inpatient rheumatology care should thus be geared more towards improving treatment outcomes, documenting these quality improvements, and establishing methods for optimizing the quality of clinical process (Harrington 2008). Additional evaluations, for example, of mental health of patients by means of PROs, as has been done in this study, may help to improve treatment outcomes (Ahles et al., 2006).

As could be shown in this article, patient reported outcomes regarding pain, mental state and complaints do improve under complex treatment conditions. Ongoing measurement of pain intensity by patients themselves is considered the gold standard of pain measurement (Turk et al., 2008). Patient information collected via VAS allows the treatment team to respond quickly and adjust the treatment. Additional monitoring of complaints, for example, by means of the Zerssen complaint score, gives the treatment team a more accurate picture of the patient status, and can be used as a basis for individual patient care and symptom control (Zerssen & Petermann, 2011). In addition, integrating PROs into multimodal treatment concepts also provides the opportunity to benchmark against other hospitals, or health systems of other countries (Van der Wees et al., 2014, Cella et al., 2015).

## 6. Conclusion

In Germany, MRT is offered by specialized clinics for the treatment of acute exacerbations of complex rheumatic diseases. Due to its legally prescribed content, it offers an intensive treatment option with longer length of stay in hospital, the added value of multiple treatments, and patient-centred outcome measurement. Specialized hospitals are granted incentives based on the quality of structure, process, and outcome specified by MRT.

Remuneration systems can be improved by incorporating P4P aspects, also regarding complex treatments. Based on patient reported outcomes, the latter have a positive impact on perceived pain, mental state and complaints. Though measured twice and with very good response rates, data using PROMs was only taken during hospitalisation, so there is no information regarding later effects. Moreover, in addition to structures and processes in hospital care, there are other factors that influence the quality of patient care which may be highly individualistic. Long-term evaluations of multimodal treatment concepts for chronically ill patients are called for.

## Acknowledgments

The researchers appreciate all feedback from the anonymous reviewers and thank all those who participated in the study and helped to facilitate the research process.

## Competing Interests Statement

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*. Washington DC: American Psychiatric Association.
- Anyfanti, P., Triantafyllou, A., Panagopoulos, P., Triantafyllou, G., Pyrpassopoulou, A., Chatzimichailidou, S., ... & Douma, S. (2016). Predictors of impaired quality of life in patients with rheumatic diseases. *Clinical rheumatology*, 35(7), 1705-1711. <https://doi.org/10.1007/s10067-015-3155-z>
- Ahles, T. A., Wasson, J. H., Seville, J. L., Johnson, D. J., Cole, B. F., Hanscom, B., ... McKinstry, E. (2006). A controlled trial of methods for managing pain in primary care patients with or without co-occurring psychosocial problems. *The Annals of Family Medicine*, 4(4), 341-350. <https://doi.org/10.1370/afm.527>
- Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S., & Guthrie, B. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*, 380(9836), 37-43. [https://doi.org/10.1016/S0140-6736\(12\)60240-2](https://doi.org/10.1016/S0140-6736(12)60240-2)
- Black, N. (2013). Patient reported outcome measures could help transform healthcare. *Bmj*, 346, f167. <https://doi.org/10.1136/bmj.f167>
- Bombardier, C., & Mian, S. (2013). Quality indicators in rheumatoid arthritis care: using measurement to promote quality improvement. *Annals of the rheumatic diseases*, 72(suppl 2), ii128-ii131. <https://doi.org/10.1136/annrheumdis-2012-202259>
- Bond, A., & Lader, M. (1974). The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*, 47(3), 211-218. <https://doi.org/10.1111/j.2044-8341.1974.tb02285.x>

- Boyd, C. M., Darer, J., Boulton, C., Fried, L. P., Boulton, L., & Wu, A. W. (2005). Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *Jama*, 294(6), 716-724. <https://doi.org/10.1001/jama.294.6.716>
- Burckhardt, C. S., & Jones, K. D. (2003). Adult measures of pain: the McGill pain questionnaire (MPQ), rheumatoid arthritis pain scale (RAPS), short-form mcgill pain questionnaire (SF-MPQ), verbal descriptive scale (VDS), visual analog scale (VAS), and west haven-yale multidisciplinary pain inventory (WHYMPI). *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 49(S5), S96-S104. <https://doi.org/10.1002/art.11440>
- Castrejon, I., & Pincus, T. (2012). Patient self-report outcomes to guide a treat-to-target strategy in clinical trials and usual clinical care of rheumatoid arthritis. *Clinical and Experimental Rheumatology-Incl Supplements*, 30(4), S50.
- Cella, D. F., Hahn, E. A., Jensen, S. E., Butt, Z., Nowinski, C. J., Rothrock, N., & Lohr, K. N. (2015). *Patient-reported outcomes in performance measurement* (Vol. 97). Research Triangle Park, NC: RTI Press.
- Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., ... & Cook, K. (2010). Initial adult health item banks and first wave testing of the patient-reported outcomes measurement information system (PROMIS™) network: 2005–2008. *Journal of Clinical Epidemiology*, 63(11), 1179. <https://doi.org/10.1016/j.jclinepi.2010.04.011>
- Chin, R., & Lee, B. Y. (2008). *Economics and patient reported outcomes, Principles and practice of clinical trial medicine*. London, Amsterdam, Burlington, San Diego: Elsevier Inc; 2008. pp. 145–66. <https://doi.org/10.1016/B978-0-12-373695-6.00008-9>
- Cross, M., Smith, E., Hoy, D., Carmona, L., Wolfe, F., Vos, T., ... & Buchbinder, R. (2014). The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Annals of the rheumatic diseases*, 73(7), 1316-1322. <https://doi.org/10.1136/annrheumdis-2013-204627>
- Dawson, J., Doll, H., Fitzpatrick, R., Jenkinson, C., & Carr, A. J. (2010). The routine use of patient reported outcome measures in healthcare settings. *Bmj*, 340. <https://doi.org/10.1136/bmj.c186>
- DiPiero, A., & Sanders, D. G. (2005). Condition based payment: improving care of chronic illness. *Bmj*, 330(7492), 654-657. <https://doi.org/10.1136/bmj.330.7492.654>
- Donabedian, A. (1966). Evaluating the quality of medical care. *Milbank Mem Fund Q.*, 44, 166–206. <https://doi.org/10.2307/3348969>
- Downie, W. W., Leatham, P. A., Rhind, V. M., Wright, V., Branco, J. A., & Anderson, J. A. (1978). Studies with pain rating scales. *Annals of the rheumatic diseases*, 37(4), 378-381. <https://doi.org/10.1136/ard.37.4.378>
- Fahrenberg, J., Hampel, R., & Selg, H. (2001). *FPI-R. Das Freiburger Persönlichkeitsinventar*, 7.
- Fautrel, B., Alten, R., Kirkham, B., De La Torre, I., Durand, F., Barry, J., ... Taylor, P. C. (2018). Call for action: how to improve use of patient-reported outcomes to guide clinical decision making in rheumatoid arthritis. *Rheumatology international*, 38(6), 935-947. <https://doi.org/10.1007/s00296-018-4005-5>
- Flynn, D., Van Schaik, P., & Van Wersch, A. (2004). A comparison of multi-item likert and visual analogue scales for the assessment of transactionally defined coping function1. *European Journal of Psychological Assessment*, 20(1), 49-58. <https://doi.org/10.1027/1015-5759.20.1.49>
- Food and Drug Administration. (2009). *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. 2009. Rockville, Food and Drug Administration.
- Frettlöh, J., Maier, C., Gockel, H., & Hüppe, M. (2003). Validität des Mainzer Stadienmodells der Schmerzchronifizierung bei unterschiedlichen Schmerzdiagnosen. *Der Schmerz*, 17(4), 240-251. <https://doi.org/10.1007/s00482-003-0227-9>
- Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of general internal medicine*, 22(11), 1596-1602. <https://doi.org/10.1007/s11606-007-0333-y>
- Gift, A. G. (1989). Visual Analogue Scales: Measurement of a subjective phenomena. *Nursing Research*, 30, 268-288. <https://doi.org/10.1097/00006199-198909000-00006>
- Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". *Diagnostica*, 50(4), 171-181.

- <https://doi.org/10.1026/0012-1924.50.4.171>
- Greenhalgh, J., Dalkin, S., Gooding, K., Gibbons, E., Wright, J., Meads, D., ... & Pawson, R. (2017). Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Health Services and Delivery Research*, 5(2), 1-280. <https://doi.org/10.3310/hsdr05020>
- Harrington, J. T. (2008). Quality of care in rheumatic diseases: performance measures and improvement. *Current opinion in rheumatology*, 20(2), 153-158. <https://doi.org/10.1097/BOR.0b013e3282f4b200>
- Harrison, M., Milbers, K., Mihic, T., & Anis, A. H. (2016). Incentives in Rheumatology: the Potential Contribution of Physician Responses to Financial Incentives, Public Reporting, and Treatment Guidelines to Health Care Sustainability. *Current rheumatology reports*, 18(7), 42. <https://doi.org/10.1007/s11926-016-0596-6>
- Haywood, K. L. (2006). Patient-reported outcome I: measuring what matters in musculoskeletal care. *Musculoskeletal Care*, 4(4), 187-203. <https://doi.org/10.1002/msc.94>
- Huskisson, E. C. (1974). Measurement of pain. *The lancet*, 304(7889), 1127-1131. [https://doi.org/10.1016/S0140-6736\(74\)90884-8](https://doi.org/10.1016/S0140-6736(74)90884-8)
- Jha, A. K., Joynt, K. E., Orav, E. J., & Epstein, A. M. (2012). The long-term effect of premier pay for performance on patient outcomes. *New England Journal of Medicine*, 366(17), 1606-1615. <https://doi.org/10.1056/NEJMsa1112351>
- Klemm, P., Hudowenz, O., Asendorf, T., Müller-Ladner, U., Lange, U., & Tarner, I. H. (2019). Multimodale rheumatologische Komplexbehandlung bei rheumatoider Arthritis—eine monozentrische Retrospektivanalyse. *Manuelle Medizin*, 57(3), 152-157. <https://doi.org/10.1007/s00337-019-0554-2>
- Kondo, K. K., Damberg, C. L., Mendelson, A., Motu' apuaka, M., Freeman, M., O' Neil, M., ... & Kansagara, D. (2016). Implementation processes and pay for performance in healthcare: a systematic review. *Journal of general internal medicine*, 31(1), 61-69. <https://doi.org/10.1007/s11606-015-3567-0>
- Kotronoulas, G., Kearney, N., Maguire, R., Harrow, A., Di Domenico, D., Croy, S., & MacGillivray, S. (2014). What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. *Journal of clinical oncology*, 32(14), 1480-1510. <https://doi.org/10.1200/JCO.2013.53.5948>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. Validity of a brief depression severity measure. *J Gen Intern Med*, 16, 606-13. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Levin-Scherz, J., DeVita, N., & Timbie, J. (2006). Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS® measures in an integrated delivery network. *Medical Care Research and Review*, 63(1\_suppl), 14S-28S. <https://doi.org/10.1177/1077558705284057>
- Mendelson, A., Kondo, K., Damberg, C., Low, A., Motu'apuaka, M., Freeman, M., ... & Kansagara, D. (2017). The effects of pay-for-performance programs on health, health care use, and processes of care: a systematic review. *Annals of Internal Medicine*, 166(5), 341-353. <https://doi.org/10.7326/M16-1881>
- Nikiphorou, E., Nurmohamed, M. T., & Szekanecz, Z. (2018). Comorbidity Burden in Rheumatic Diseases. *Frontiers in medicine*, 5, 197. <https://doi.org/10.3389/fmed.2018.00197>
- Richard, M. (2010). *Pay for performance (P4P) programs in health services: what is the evidence*. World Health Report. Background Paper, 31.
- Roberts, E. T., Zaslavsky, A. M., & McWilliams, J. M. (2018). The value-based payment modifier: program outcomes and implications for disparities. *Annals of internal medicine*, 168(4), 255-265. <https://doi.org/10.7326/M17-1740>
- Romeyke, T., & Stummer, H. (2014). Multi-modal pain therapy of fibromyalgia syndrome with integration of systemic whole-body hyperthermia—effects on pain intensity and mental state: a non-randomised controlled study. *Journal of Musculoskeletal Pain*, 22(4), 341-355. <https://doi.org/10.3109/10582452.2014.949336>
- Romeyke, T., Noehammer, E., & Stummer, H. (2018). Interdisciplinary assessment-oriented treatment of fibromyalgia: a case report. *Integrative medicine research*, 7(2), 200-205. <https://doi.org/10.1016/j.imr.2018.02.002>
- Rotenstein, L. S., Huckman, R. S., & Wagle, N. W. (2017). Making patients and doctors happier—the potential of patient-reported outcomes. *New England Journal of Medicine*, 377(14), 1309-1312.

- <https://doi.org/10.1056/NEJMp1707537>
- Ryan, A. M., & Doran, T. (2012). The effect of improving processes of care on patient outcomes: evidence from the United Kingdom's quality and outcomes framework. *Medical care*, 191-199. <https://doi.org/10.1097/MLR.0b013e318244e6b5>
- Ryan, A., & Blustein, J. (2012). Making the best of hospital pay for performance. *New England Journal of Medicine*, 366(17), 1557-1559. <https://doi.org/10.1056/NEJMp1202563>
- Ryan, A. M., Blustein, J., Doran, T., D. Michelow, M., & Casalino, L. P. (2012). The effect of phase 2 of the premier hospital quality incentive demonstration on incentive payments to hospitals caring for disadvantaged patients. *Health services research*, 47(4), 1418-1436. <https://doi.org/10.1111/j.1475-6773.2012.01393.x>
- Schoen, C. & Osborn, R. (2010). *Commonwealth Fund international health policy survey*. Retrieved from <http://www.commonwealthfund.org/Surveys/2010/Nov/2010-International-Survey.aspx>
- Scott, A., Schurer, S., Jensen, P. H., & Sivey, P. (2009). The effects of an incentive program on quality of care in diabetes management. *Health economics*, 18(9), 1091-1108. <https://doi.org/10.1002/hec.1536>
- Siegel, S., & Castellan, N. J. (1956). *Nonparametric statistics for the behavioral sciences (Vol. 7)*. New York: McGraw-hill.
- Simpson, C. R., Hannaford, P. C., McGovern, M., Taylor, M. W., Green, P. N., Lefevre, K., & Williams, D. J. (2007). Are different groups of patients with stroke more likely to be excluded from the new UK general medical services contract? A cross-sectional retrospective analysis of a large primary care population. *BMC family practice*, 8(1), 56. <https://doi.org/10.1186/1471-2296-8-56>
- Sloan, J. A., Halyard, M. Y., Frost, M. H., Dueck, A. C., Teschendorf, B., Rothman, M. L., & Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. (2007). The Mayo Clinic manuscript series relative to the discussion, dissemination, and operationalization of the Food and Drug Administration guidance on patient-reported outcomes. *Value in health*, 10, S59-S63. <https://doi.org/10.1111/j.1524-4733.2007.00267.x>
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*, 282(18), 1737-1744. <https://doi.org/10.1001/jama.282.18.1737>
- Stafford, L., Berk, M., & Jackson, H. J. (2007). Validity of the Hospital Anxiety and Depression Scale and Patient Health Questionnaire-9 to screen for depression in patients with coronary artery disease. *General hospital psychiatry*, 29(5), 417-424. <https://doi.org/10.1016/j.genhosppsych.2007.06.005>
- Turk, D. C., Dworkin, R. H., Revicki, D., Harding, G., Burke, L. B., Cella, D., ... & Max, M. B. (2008). Identifying important outcome domains for chronic pain clinical trials: an IMMPACT survey of people with pain. *PAIN®*, 137(2), 276-285. <https://doi.org/10.1016/j.pain.2007.09.002>
- US Preventive Services Task Force, 2002
- Valderas, J., Kotzeva, A., Espallargues, M., Guyatt, G., Ferrans, C. E., Halyard, M. Y., ... & Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Quality of life research*, 17(2), 179-193. <https://doi.org/10.1007/s11136-007-9295-0>
- Van Der Heijde, D., Daikh, D. I., Betteridge, N., Burmester, G. R., Hassett, A. L., Matteson, E. L., ... & Lakhanpal, S. (2018). Common language description of the term rheumatic and musculoskeletal diseases (RMDs) for use in communication with the lay public, healthcare providers and other stakeholders endorsed by the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR). *Annals of the rheumatic diseases*, 77(6), 829-832. <https://doi.org/10.1136/annrheumdis-2017-212565>
- Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. B., & Sermeus, W. (2010). Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC health services research*, 10(1), 247. <https://doi.org/10.1186/1472-6963-10-247>
- Van der Wees, P. J., Nijhuis-Van Der Sanden, M. W., Ayanian, J. Z., Black, N., Westert, G. P., & Schneider, E. C. (2014). Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: Views of experts from 3 countries. *The Milbank Quarterly*, 92(4), 754-775. <https://doi.org/10.1111/1468-0009.12091>
- Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., ... & Abraham, J. (2012). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for



the Global Burden of Disease Study 2010. *The lancet*, 380(9859), 2163-2196.

Wurmthaler, C., Gerbershagen, H.U., Dietz, G., Korb, J., Nilges, P., & Schillig, S. (1996). Chronifizierung und psychologische Merkmale -die Beziehung zwischen Chronifizierungsstadien bei Schmerz und psychophysischem Befinden, Behinderung und familiären Merkmalen. *Z Gesundheitspsych* 4, 113-136.

Zerssen, V. D., & Koeller, D. M. (1976). *Klinische Selbstbeurteilungs-Skalen (KSB-S) aus dem Muenchener Psychiatrischen Informations-System*. Die Befindlichkeitsskala

Zerssen, V. D., & Petermann, F. (2011). *B-LR – Beschwerden-Liste – Revidierte Fassung*. Hogrefe, Göttingen.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).