

Aus dem Institut für Gesundheits- und Rehabilitationswissenschaften
der Ludwig-Maximilians-Universität München
Ehemaliger Vorstand: Prof. Dr. med. Gerold Stucki
Kommissarischer Vorstand: Prof. Dr. med. Dennis Nowak

Predictors of willingness to pay for health benefits:
an exploratory analysis of
willingness to pay for small to moderate health effects among
musculoskeletal and psychosomatic patients
undergoing rehabilitation in Germany

Dissertation
zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Doris Gabriele Gerstner
aus
Rosenheim
2011

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	Prof. Dr. med. Gerold Stucki
Mitberichterstatter:	Prof. Dr. Jürgen Stausberg Priv. Doz. Dr. Johanna Anneser
Mitbetreuung:	Priv. Doz. Dr. Alarcos Cieza
Dekan:	Prof. Dr. med. Dr. h. c. M. Reiser, FACR, FRCR
Tag der mündlichen Prüfung:	12.01.2011

First and foremost, I thank Professor Gerold Stucki and Priv. Doz. Dr. Alarcos Cieza for their encouragement, advice and vision throughout the process of this thesis. A special thanks to Mirjam Brach for her help and for the excellent working conditions provided at the Department of Physical Medicine and Rehabilitation.

I would also like to thank my colleagues for creating an encouraging and stimulating working atmosphere, especially Dr. Marita Stier-Jarmer as we shared an office throughout this time, thank you for being there.

Lastly, I am grateful to have been blessed with a loving and supporting family. My love goes to my husband Georg and my two daughters Anna-Katharina and Sophie-Marie, who always showed interest and support in the progress that I was making.

Contents

1	Background	3
1.1	The willingness to pay measure	3
1.2	Willingness to pay and respondents` characteristics: going beyond income .	7
1.3	Predictors of willingness to pay using small to moderate health effects of rehabilitation interventions as a case in point.....	9
2	Research objectives	15
3	A systematic review of predictors of willingness to pay for changes in pain and related symptoms	16
3.1	Objective	16
3.2	Methods.....	16
3.3	Results	19
3.3.1	Predictors linked to the ICF	20
3.3.2	Predictors not linked to the ICF	25
3.4	Discussion	27
4	An exploratory analysis of predictor candidates of willingness to pay for small to moderate health effects of rehabilitation interventions.....	40
4.1	Objective	40
4.2	Methods.....	40
4.2.1	Study population	40
4.2.2	Data	41
4.2.3	Willingness to pay measurement	42
4.2.4	Predictors.....	42
4.2.5	Statistical analysis.....	44
4.3	Results	47
4.3.1	LASSO	47
4.3.2	Stepwise-built GLM.....	48
4.4	Discussion	48
5	Conclusion.....	58
6	Summary	62
7	Zusammenfassung	67
8	References	73

9	Appendix	85
	Appendix 1: Search strategy	85
	Appendix 2: Explanations for significant predictors	86
	Appendix 3: Content of scales and meaning of scores	91
10	Curriculum Vitae	94

1 Background

1.1 The willingness to pay measure

Economic issues involved in the evaluation of health and health care interventions are recognised in many countries to be of great importance. In the health arena, policy makers and program managers are constantly faced with economic decisions; one of the most pertinent being how to spend a limited budget and have the biggest positive impact on health. The technique of economic evaluation can contribute to these decisions by providing information on the costs and benefits of alternative interventions. The ultimate aim of an economic evaluation is to determine whether the benefits of an intervention exceed its costs [1].

Historically, cost-effectiveness and cost-utility analysis have been the most widely used types of economic evaluation of medical procedures and technologies. More recently, there is growing interest in cost benefit analysis and the use of the concept 'willingness to pay' (WTP) for measuring health benefits [2, 3]. A cost benefit analysis values costs and benefits in monetary terms [4]. Only if the benefits of a health program exceed its costs which can be seen by a positive net benefit when subtracting costs from benefits, the program constitutes a true welfare improvement. By nature, cost benefit analysis allows to directly draw conclusions about the allocative efficiency of a health care program [5].

The most common measure of benefit in monetary terms is willingness to pay (WTP). This measure derived from welfare economic theory relies on the basic premise that the maximum amount of money individuals are willing to pay for a health gain is an indicator of the value of that health gain to them [6, 7]. The WTP measure

is justified as representing the preferences of individuals and is allowed to depend on health consequences and other individual characteristics, as well as on characteristics of the good being valued [8]. WTP estimates can be derived using different methods such as the elicitation of revealed preferences or stated preferences. Revealed preference approaches which are preferred by economists infer the value of a good or service from market transactions [9]. However, when information on WTP cannot be based on observable behavior because the good or service is not provided in a market, as it is often the case in the healthcare sector, then a need for using survey techniques to elicit stated WTP arises. Stated preference methods include techniques such as contingent valuation or conjoint analysis [10]. Although adopted more extensively in environmental economics, the elicitation of WTP in a contingent valuation survey to value health outcomes, has increased significantly over the last decade [11-18].

Contingent valuation involves asking individuals directly in a hypothetical survey the maximum amount they are willing to pay to obtain a specific health improvement. In other words, the contingent valuation relies on the stated intentions of individual's WTP for a given health change, contingent on a market that is hypothetical, hence the name contingent valuation [19]. The ideal in a contingent valuation survey is to get the respondents to make hypothetical choices in the same way they would if faced with an actual decision situation.

However, eliciting the WTP based on contingent valuation surveys (hereafter referred to as "WTP") is challenging. By its very nature the WTP represents hypothetical answers which may correlate poorly with what the individual would actually pay. This difference between hypothetical (stated) and actual (revealed)

values is termed hypothetical bias [20]. Moreover, since it is not common in many countries, providing its residents with comprehensive health insurance coverage, to pay directly for healthcare, individuals may feel unfamiliar in their decisions and may not have an adequate basis for articulating their true value. Recent literature on WTP therefore emphasizes the importance of creating clearly defined and realistic scenarios [13]. With regard to the possibility of a hypothetical bias invalidating the results, it is essential to evaluate the validity of WTP responses [21].

Concerning the hypothetical nature the gold standard for proving the validity of WTP is the comparison of stated WTP with actual payments [4]. However, such comparisons are difficult to conduct in the absence of a real market as is the case for non market goods such as most health programs. Hence, it is difficult to establish criterion validity in the context of health care.

Targeting theoretical validity is an alternative and usual way to address the validity of WTP responses. It involves measuring whether the WTP data conform to hypotheses as predicted by theory or as related to existing empirical evidence. Theoretically, the WTP approach assumes that the valuation depends on both the characteristics of the good being valued and the characteristics of the individual. With regard to the “good”, economic theory suggests that the WTP should be sensitive to the amount of the good, or service supplied, which is tested in scope tests. Typically, WTP should increase with higher benefits. This issue is of utmost importance in the evaluation of health care programs which are targeted to improve outcomes. Therefore, given the extent that the WTP method is used to value improvements in outcomes, it is crucial that valuations are sensitive to the size of these outcomes [22]. However, the evidence of sensitivity to health changes varied for WTP values. Whilst

a few studies have found strong indications of scope sensitivity, others yielded evidence of the insensitivity of WTP to the size of health improvements [14]. Scope tests should therefore be part of every WTP study, but are addressed only in a few studies [22].

Concerning respondents' characteristics, economic theory assumes that WTP should vary with ability to pay which is commonly inferred from income information [4, 16]. Lack of a positive income effect is commonly explained as an indication that respondents did not seriously account for their budget constraints when making hypothetical choices [23]. The influence of income on WTP is not without criticism because it may result in health policy decisions in favour of the rich if their preferences differ from the poor [24, 25]. In cases where richer people tend to prefer an option and are willing to pay more, the WTP unadjusted for income will skew resource allocation to the preferences stated by the wealthy. However, statistical techniques can be used to control for the effect of income and to estimate an income adjusted WTP [17]. This may help to reduce scepticism about using the WTP method to value health benefits and may divert attention from inequalities due to its income dependency. As pointed out in a review, over time, there has been a different handling of the income variable as an equity issue, and the influence of income on WTP is currently considered as a minimum validity check [15]. Many WTP studies have assessed the theoretical validity through considering the relationship between WTP and ability to pay. Indeed, income has been shown to be an important economic determinant of WTP in most studies (see e.g. [26-30]).

1.2 Willingness to pay and respondents` characteristics: going beyond income

Despite empirical evidence for an income dependency of WTP, there remains considerable uncertainty concerning variations in WTP responses. Income cannot fully explain how individuals value potential health benefits. For example, income was the best predicting factor, but explained only 14% of WTP in a study on a hypothetical cure of chronic obstructive pulmonary disease [31]. Other personal factors did not explain additional variance. Such a lack of explanatory power of the regression model has been shown as well in another WTP study about chronic obstructive pulmonary disease [32] and in several studies estimating the WTP for treatment effects of other health conditions. The explained variance of the WTP for antihypertensive therapy, for example, ranged from 12% to 14% and was provided by significant (income and health benefit) and non significant variables [33]. Or, as pointed out in a study to elicit the economic value of an improved malaria treatment program, income was the variable most significant in a model with an explanatory power of 7% [34]. Last but not least, the WTP for an imaginary treatment of epilepsy showed a significant association only with income which explained 13% of the variability in WTP [35].

Apart from income, the WTP literature suggests existing additional respondents` characteristics often not directly derived from standard economic theory may explain variation in WTP. Published guidelines for health care WTP studies recommend investigating the influence of respondents` characteristics on WTP values as an ongoing validation of data from such studies [18, 36]. The importance of testing theoretical validity in WTP studies was emphasized when prominent

economists assembled by the National Oceanic and Atmospheric Administration (NOAA) panel published a set of general guidelines on the design and implementation of contingent valuation studies for environmental damage assessment in light of the Exxon Valdez oil spill in Alaska [37]. Along with their specific guidelines the NOAA panel issued the recommendation to break down WTP by introducing a variety of respondents' characteristics such as income, interest, and attitudes to better understand the WTP responses. Although developed in the environmental field, the NOAA recommendations are thought to be applicable in the health care sector [18]. Hence, as a measure of validity, it is standard practice for health care WTP studies to link elicited contingent values to personal characteristics. For example, socio-demographic and economic factors were regressed on WTP in a study to estimate the value that the Canadian tax-paying public puts on a new cyto-protective agent that reduces the risk of chemotherapy-induced toxicity. The results suggested that respondents with higher family income and married/common law respondents were willing to pay significantly higher amounts as those with lower income and those who were single. Significant regional differences (participants from Quebec were willing to pay more than their counterparts in Ontario) were also observed. All three predictors explained 30% of the variability of WTP [38].

Besides socio-demographic variables a variety of other individual characteristics showed significant relationships to WTP. For example, together with demographic variables the influence of visual impairment was investigated in a study to evaluate the WTP for cataract surgery in a typical, poor, rural region of south China [39]. The results demonstrated that after adjustment for income visually impaired and blind persons were more willing to pay anything for cataract surgery, but the amount that blind persons were willing to pay was significantly less than for

persons with good vision. The results also suggested a higher WTP among younger patients.

Furthermore psychosocial factors were used to examine their influence on the public` s WTP for genetic testing to assess cancer risk [40]. Being aware of genetic susceptibility, likelihood of undergoing genetic testing for cancer, and talking and seeking information about it, was directly associated with an increased chance of being willing to pay more, independent of other indirect associations such as e.g. perceived benefits (information about cancer risk, better decision making) and socioeconomic status (education, household income).

The list of persons` characteristics influencing the WTP can be extended due to a large number of studies investigating the drivers of WTP in the context of health care. Drawing conclusions based upon the evidence of predictors derived from these studies, however, is difficult. To date, there is no systematic overview of factors affecting the WTP beyond income, and consequently there is a lack of knowledge about the impact of persons` characteristics on the WTP for health benefits.

1.3 Predictors of willingness to pay using small to moderate health effects of rehabilitation interventions as a case in point

In health care the WTP based on contingent valuation has been found to be a valuable approach to assess different conditions and programs, and different sizes of health gains such as a cure or partial relief. Moreover, WTP appears to be a promising technique when measuring small but meaningful effects of interventions.

Incremental improvements in the efficacy of treatments are likely to be the rule for many current treatment regimens and one typical example is the rehabilitation of patients suffering from chronic conditions [41-45]. Based on the International Classification of Functioning, Disability, and Health (ICF) of the World Health Organization [46], rehabilitation can be defined as a health strategy that “aims to enable people with health conditions experiencing or likely to experience disability to achieve optimal functioning in interaction with the environment” [47]. In this context, the WTP based on contingent valuation has been shown to be a feasible and valid approach. In a study by Brach M, Gerstner DG, Stucki G (Usefulness of WTP to evaluate small to moderate clinical effects in rehabilitation, to be provided for publication in 2010) it could be demonstrated that the WTP of patients with five different types of health conditions (osteoarthritis, osteoporosis, back pain, pain disorder, depression) depended on income and showed a tendency to be sensitive to health gains, the so-called scope effect. These findings are consistent with the theoretical framework of WTP which assumes that WTP should increase with respondents` income and with the size of the good being valued [4].

Proper use of methods of cost-benefit analysis in specific settings such as for the evaluation of small to moderate health effects in rehabilitation relies on an in-depth understanding of factors influencing the results of such analyses. As previously stated in this thesis, research from other studies developing methods for cost-benefit analysis suggests that in addition to income there are other respondents` characteristics that drive respondents` WTP for health benefits. Obviously, these factors need to be explored when aiming to understand the drivers of WTP for small to moderate health effects of rehabilitation interventions. It is therefore, important to investigate the factors that have been shown to affect the WTP for health outcomes

systematically. Most importantly, we need to understand their relevance from a comprehensive perspective. So far, no study has evaluated all current putative predictors of WTP simultaneously, so their relative contribution to WTP is as yet not understood.

Although many WTP studies have investigated predictors related to WTP for health effects gained from treating the same condition, no study has been conducted to determine whether predictors vary across different health conditions. From outcome research it is well known that patients' problems vary greatly across health conditions. One can therefore expect the factors that explain the WTP to vary, to a certain extent, across health conditions. In order to explore this expectation five chronic health conditions (osteoarthritis, osteoporosis, back pain, pain disorder, and depression) which differ with respect to the spectrum of their clinical presentation and appropriate intervention strategies were used.

Osteoarthritis (OA) is the most common joint disease in the world and prevalence is likely to increase because of the aging worldwide population [48, 49]. The principal morphological characteristic of OA is a slowly developing degenerative breakdown of cartilage. Relatively little is known of the natural history of OA. Symptoms and the radiographic appearance of OA do not always coincide, making the early diagnosis and monitoring of the disease difficult. From a clinical perspective, evaluations of OA patients and decisions for treatment strategy have to consider not only radiological findings but also patients' symptoms. Pain or discomfort, limitations of activity and reduced participation in daily activities are key concerns associated with OA [50]. Because of the frequent disability that accompanies OA in the knee and hip and its prevalence, OA accounts for more

problems with activities such as climbing stairs and walking than any other health condition [51].

Recently, the WHO has defined Osteoporosis (OP) as a disease that is characterized by low bone mass and micro-architectural deterioration of bone tissue with a resulting increase in bone fragility and susceptibility to fracture [52-54]. OP primarily affects older white women. According to the WHO Technical Report Series No. 843 [53], about 30% of postmenopausal women are estimated to have OP. It is difficult to diagnose early because symptoms (e.g. back pain, loss of height, spinal deformity) often do not occur until late in life, after considerable loss of bone mass has already occurred. Disability mainly occurs after a fracture. There may be little or no impact in the pre-fracture stage, however, most of the burden of OP relates to fracture. Pain with loss of function is the major outcome of any fracture [55]. Activity limitations and restrictions in participation mostly come from fractures or fracture-related complications; many may lose their independence after multiple fractures [56].

Back pain is a common phenomenon in Western societies, and the most common cause of disability in people under 45 years of age [57]. Back pain, and low back pain (LBP) in particular, affects 60% and 80% of people at some point in their lives [58, 59]. The majority (approximately 95% of cases) of acute LBP is non-specific; serious conditions are rare. Common findings e.g. osteoarthritis, lumbar spondylosis and spinal canal stenosis also occur in asymptomatic people; hence, such conditions may not be the cause of the pain [60]. Non-specific back pain is thus a major problem for diagnosis and treatment. LBP is the most common physical condition for which patients visit their doctor [61]. Many countries have produced

guidelines for the management of LBP. Although the precise details of the national guidelines vary they all have as components the early and gradual activation of patients, the avoidance of bed rest, and the acknowledgement of the role of psychosocial factors as risk factors for chronicity [62]. Besides its influence on the progression from acute to chronic pain, psychosocial factors are believed to be important for the onset of LBP and for the response to treatment [63]. Now, there is a shift “from thinking about back pain as a biomedical injury towards viewing LBP as a multifactorial biopsychosocial pain syndrome” [64, 65].

Perhaps the most common psychiatric disorder diagnosis that chronic pain patients receive is pain disorder (PD) [66]. PD is classified in the DSM-IV [67] in the chapter of somatoform disorders which are characterized by the presentation of physical symptoms despite negative findings of organic illness (hence, the term somatoform). The diagnostic criteria for pain disorder in DSM-IV specifies that pain is the predominant clinical feature, causes significant distress or functional impairment, and that psychological characteristics are judged to play an important role in the onset, severity, exacerbation, or maintenance of the pain. Thus, the criterion of absent / insufficient organic findings as defined in the DSM-III [68] was removed from the DSM-IV. Furthermore, the DSM-IV no longer stipulates an etiologic relationship with psychological factors. Instead, the DSM-IV requires only that psychological factors play an important role [66]. Due to the changes in diagnostic criteria of pain, and the subjectivity required to make these diagnoses, prevalence rates varied considerable. A study referring to the DSM-IV reported that pain disorder was prevalent in 8.1% of the general German population. DSM-IV PD was more frequent in women (men: 4.3%; women: 11.4%) [69].

Depression is a highly prevalent condition that results in substantial patient suffering, family distress and conflict, and a significantly increased risk of suicide. Global Burden of Disease 2000 (GBD) analysis estimates that 5.8% of men and 9.5% of women will experience a depressive episode in a 12-month period [70]. The GBD 2000 analysis also shows that unipolar depressive disorders are ranked as the fourth leading cause of burden among all diseases. Depression can affect individuals at any stage of the lifespan, although the incidence is highest in middle age. Depression is a chronic disorder causing a very high level of disease burden. The symptoms of depression can be cognitive (e.g. reduced concentration), behavioral (e.g. social withdrawal), and physical (e.g. bodily pain). In fact, limitations and restrictions in activities and participation may be most relevant to patients with depression [71].

To summarize, all five chronic health conditions are common disorders which produce an enormous burden of disease and cause a major public health problem. Each chronic condition has its typical spectrum of problems in functioning. It can therefore be expected that the predictors of WTP may to some extent differ across these health conditions.

2 Research objectives

The overall objective of this doctoral thesis was to explore respondents' characteristics beyond income and scope predicting the WTP for health benefits from contingent valuation studies.

The specific aims were 1) to identify relevant predictors from other studies investigating the WTP for changes in pain and related symptoms, and 2) to examine putative predictors of WTP for small to moderate health effects of rehabilitation interventions.

With respect to these specific aims the doctoral thesis was subdivided into two parts. The first part presents predictor candidates of WTP based on a systematic review focusing on WTP for changes in pain and related symptoms. The review focuses on this issue because pain relief was the main outcome to be valued in the WTP study to which the subsequent predictor analysis refers. The second part presents an exploratory analysis of predictors of WTP for small to moderate health effects among musculoskeletal and psychosomatic patients undergoing rehabilitation. A comprehensive set of predictor candidates including aspects of functioning and personal factors was used as predictor candidates and the influence on WTP was compared across five different health conditions (osteoarthritis, osteoporosis, back pain, pain disorder, and depression). Each of these parts contains a respective discussion section referring to the specific results.

3 A systematic review of predictors of willingness to pay for changes in pain and related symptoms

3.1 Objective

The objective of this part of the doctoral thesis was to identify predictors of WTP beyond income, focusing on WTP for changes in pain and related symptoms. For this purpose a systematic review of published WTP studies based on contingent valuation was carried out.

3.2 Methods

Scientific literature was reviewed to identify studies that investigated predictors of WTP for changes in pain and related symptoms. A computer-assisted search of different databases was performed, including WebSpirs MEDLINE (R) from 1966 to March week 1-2/2008, WebSpirs EMBASE (R) from 1989 to January 2008, WebSpirs PsycInfo from 1806 to March week 1 / 2008, and EconLit from 1969 to March 2008. The full search strategy can be seen in Appendix 1.

Included studies had to have elicited the WTP for changes in pain and related symptoms either as partial or complete relief, and provide predictors of WTP in addition to economic factors like income. This review excluded publications of WTP studies: (i) based on other methods than contingent valuation (e.g. discrete choice experiments which ask individuals to choose among programs with different attributes and provide WTP information when price is treated as an attribute); (ii) without a multivariate analysis of potential predictors controlled for income; (iii) in

languages others than English; and (iv) undertaken as a systematic review or meta-analysis.

The abstracts of all the citations were retrieved and subjected to the exclusion criteria. The author was responsible for the entire selection. A second reviewer (MB) checked the selection procedure by screening a random sample (n=200) of all articles. Full text copies were then retrieved of all included articles or for those cases where inclusion was uncertain.

In cases where the principal interest was to determine the degree of relationship between WTP and another variable measuring the same construct, as is the case for testing convergent validity, then such a variable was not considered as a predictor. An example is the comparison with other methods for estimating WTP such as, e.g. the averting-behavior approach [72].

The judgement as to whether a WTP study had addressed changes in pain or related symptoms was based not only on the actual WTP question but also on the scenario description in accordance with recommendations in the WTP literature [15]. Both sources allowed the appraisal of which health benefits respondents were valuing. Despite recommendation not every study provided the necessary information in detail. Therefore, if the WTP question/scenario did not explicitly refer to pain and related symptoms (e.g. in cases where respondents were asked “to get rid of their symptoms”), or if respondents` WTP for a treatment was asked for but a scenario was not presented, the inclusion of a study then depended on whether pain and related symptoms were described in the study as an important aspect of the disease or treatment.

A standardized data extraction form was used to register the predictors analyzed, their significance, method of multivariate analysis, study objective, study population, sample size, study design (survey method and elicitation format), year of publication, and the country where the study was carried out.

The International Classification of Functioning, Disability and Health (ICF) of the World Health Organization [46] was used as a tool to classify the predictors of WTP in the health care field. This classification system is a widely accepted framework [73] and provides a common language for the description of health-related phenomena [74]. The ICF is made up of three components that address functioning and disability, namely Body Structures, Body Functions, Activities and Participation. Body Functions are defined as the physiological functions of body systems, including psychological functions. Body Structures are the anatomical parts of the body, such as organs, limbs, and their components. Activity is the individual's ability to perform a task or action, Participation refers to an individual's involvement in a real life situation. In addition to the components covering functioning and disability, the ICF contains Environmental Factors as a fourth component, describing contextual factors such as the physical, social, and attitudinal environment in which people live and lead their lives [46]. The units of the ICF classification are called categories. They are organized within a hierarchically nested structure and are denoted by unique alphanumeric codes. Each component consists of chapters (categories at the first level), each chapter consists of second level categories, and in turn they are made up of categories of the third level, and so on. For the purpose of this review, the ICF classification was used at first and second level categories. An example selected from the component Body Functions is presented in the following:

b2 Sensory functions and pain (first level category)
b280 Sensation of pain (second level category)

The extracted predictors were linked to the ICF using established linking rules [75]. For example, the pain subscale score of the Western Ontario McMaster Universities (WOMAC) has been linked to the second level ICF category b280 Sensation of pain. Personal Factors which are not currently classified by the ICF but considered a part of the contextual factors within the ICF's biopsychosocial model of human functioning and disability are the particular background of an individual's life and living situation and may include any characteristics that play a role in disability. These factors were mapped to the ICF based on self-defined categories, e.g. coping or knowledge. Predictors which lay outside the scope of the ICF classification, e.g. general health or health conditions, were presented individually.

To summarize the results, significant and non-significant predictors of WTP of multivariate models controlled for income were identified. An association with a p-value ≤ 0.05 was regarded as significant. If studies reported more than one statistical model, this review referred to the model that explained the largest variance or allowed for the best individual prediction, and included income. When results of analyses of different subgroups or different scenarios / WTP methods were reported, those results were then considered to indicate a predictor if such a predictor was found in at least one of the subgroups or scenario / WTP method groups.

3.3 Results

The electronic literature search identified 499 references without duplicates, and books published between 1966 and March 2008. Manual search revealed three

additional relevant studies. After applying the exclusion criteria a total of 21 studies remained for data extraction.

Table 1 describes the included studies [72, 76-95]. Pain as a symptom of many diseases is reflected by the distribution of studies over different health conditions such as, e.g. myalgia, multiple myeloma or migraine. The most commonly studied diseases were musculoskeletal diseases (n=9) and cardiovascular diseases (n=3). A few studies administered surveys to a general population (n=4), whereas the majority of studies focused on a patient population (n=17).

The WTP amounts were derived by different survey methods and elicitation formats. The use of interviews (n=9) and payment cards (n=9) were the dominant features of methods. Most of the reviewed studies were conducted in the USA / Canada (n=12) and Europe (n=7).

Not all studies reported the amount of variance that was explained by the predictors. Twelve studies revealed R-squares which varied from 8% to 55%, depending on model specifications.

3.3.1 Predictors linked to the ICF

Table 2 presents the distribution of predictors across the major components of the ICF. Except Body Functions and Personal Factors, which were more frequently investigated, all the other components Body Structures, Activities & Participation and Environmental Factors (except assets such as, e.g. income, which were included as control variable) were rarely addressed by the identified predictors. Predictors

referring to the ICF category b280 sensation of pain and four personal factors (age, education, gender, experiences) were assessed in at least six studies.

Body Structures

One study included predictors referring to this ICF component. It failed to find a significant relationship when connecting the T-score at femur / spine to WTP of osteoporosis patients to reduce the risk of hip fracture [95].

Body Functions

Twelve of 21 studies investigated at least one predictor referring to Body Functions. Four studies examined the influence of different aspects of functioning summarized in a single measurement score which consequently could be linked only to the component level. In one of these studies, patients with worse functioning measured with a disease specific instrument (western modification of the Japanese Orthopaedic Association / mJOA scale) were willing to pay significantly more for a cure of health problems due to their cervical spondylotic myelopathy [90].

Two of six studies including factors related to sensation of pain (b280 sensation of pain) showed a significant relationship with WTP. Subjects with headaches of unknown duration were less willing to pay high amounts for pain relief and associated symptoms of migraine attacks [87]. Two – three years after hip / knee replacement less pain (lower WOMAC pain scale) was significantly related to a higher WTP for joint replacement in patients with Osteoarthritis [81]. The association with night pain was of borderline significance ($p=0.06$) in a study to assess the strength of preferences for a treatment for patients with soft tissue rheumatism [92].

Two studies included factors related to mental functions (b152 emotional functions) and both showed significance. A study on WTP to eliminate symptoms of arthritis demonstrated that WTP increased with less depression (Centre for Epidemiological Studies Depression Scale / CES_D) in patients with rheumatoid arthritis [93]. In the other study, patients with higher preoperative anxiety (Spielberger's State-Trait Anxiety Inventory / STAIT) state levels were willing to pay more to avoid postoperative pain [94].

Higher frequency of angina symptoms or attacks (b410 heart functions) was significantly associated with higher WTP in two studies [72, 89]. In one of these studies increasing angina symptoms reached only borderline significance ($p < 0.10$) [72]. However, an increasing frequency of angina episodes in subjects who had had bypass surgery lowered the WTP significantly in that study.

Activities and Participation

One of three studies investigating the WTP of patients suffering from different forms of arthritis showed a significant relationship with limitations in activities and participation [93]. In that study, rheumatoid patients with greater restrictions in daily activities as measured by the Health Assessment Questionnaire (HAQ) and with the time required for the 50-foot walk (d450 walking) were willing to pay significantly higher WTP amounts for a cure. Subjects with more limitations caused by migraine (Migraine Disability / MIDAS score) were willing to pay more for a perfect drug [87].

Environmental Factors

One of three studies that investigated environmental factors found them to be predictive. If patients with rheumatoid arthritis were usually cared for in a public clinic

(e580 health services), then they were less willing to pay additional premiums for a private health insurance for treatment that cured their arthritis 100% [85].

Personal Factors

Research on predictors of WTP has focused primarily on socio-demographic variables. Almost all of the included studies contained at least two socio-demographic variables and six studies included five or more variables referring to this domain.

Age was significantly related to WTP in six of 17 studies. With the exception of one study [81], there was a negative relation to WTP, meaning that WTP decreased with older age [80, 85, 89, 91, 93].

Education, investigated in eight studies, predicted WTP significantly in two studies, but in opposite directions. In the study of Narbro [91], increasing WTP was related with higher education level in obese patients, and in the study of migraine sufferers the well educated were willing to pay less [87].

Employment and marital status failed to show any significance in the studies examining these factors.

Gender failed to show a significant relation to WTP in almost all of the 15 studies including this factor. Only one study showed that men of the general population were willing to pay significantly higher amounts to avoid mild pain evoked by shingles [78].

Having private health insurance increased WTP for a joint replacement [81] and a hypothetical cure for rheumatoid arthritis [85].

Number of persons / children in a household was associated with WTP in two studies. A higher number of persons decreased the WTP in a study to gain relief from overweight-related problems [91], while in another study persons having children stated a higher WTP to reduce bone fractures in multiple myeloma [83]. In the latter of the two studies, religion affected the WTP, non-Christians were willing to pay more to diminish bone fractures.

Ethnicity showed a significant relation in one of four studies exploring its influence. African Americans were willing to pay significantly less than whites for improvement in severe and mild osteoarthritis [80].

Higher social classes (a method to classify people in UK according to the occupation of the senior adult) were willing to pay less to undergo laparoscopic rather than conventional cholecystectomy [82].

Nearly half (n=10) of the 21 included studies investigated factors other than socio-demographic factors. Five studies found significant predictors of WTP:

Migraine sufferers using few mechanisms of daily coping were less likely to be willing to pay high amounts for relief [87].

Osteoporosis patients with less knowledge of osteoporosis were willing to pay more for a drug that would reduce the risk of hip fracture [95].

Prior experience of coronary artery bypass surgery lowered the WTP for percutaneous coronary interventions with a reduced risk of restenosis [86] and increased the WTP to avoid angina episodes [72] – in the latter study only in combination with lower current frequency of angina attacks. The effect of prior shingles experience decreased the WTP to move from one pain state to another in subjects of the general population [78]. Having had postoperative myalgia after previous surgery emerged as borderline significant ($p=0.07$) for the WTP to avoid postoperative muscle pain [76].

3.3.2 Predictors not linked to the ICF

As shown in Table 3, there were additional predictors investigated in the reviewed studies which were not covered by the ICF.

Two out of six studies which examined the influence of health condition showed its predictive value. In one study, a more severe angina pectoris status (unstable vs. stable) was significantly associated with a higher WTP to reduce angina pectoris attacks [89]. The other study which investigated the WTP for relief of overweight-related problems found a higher WTP with increasing severity of illness as indicated by higher body weight of obese patients [91].

One of four studies including predictors referring to general health found a significant relation to WTP. Rheumatoid arthritis patients with poorer current health, measured by the correspondent subscale of the Rand Health Perceptions Questionnaire, stated a significantly higher WTP for a complete cure [93].

Satisfaction with healthcare services or system increased WTP significantly in all three studies investigating this predictor [81, 85, 92]. In one of these studies [81], satisfaction, as measured directly was not significant, but became significant when measured indirectly by recommendation. The other two studies also measured satisfaction indirectly by asking respondents to rate their opinion of the existing health care system on a scale from poor to very good [85], or of treatment, based on a scale from very poor to very good [92].

One of three studies investigating health expenditures showed a significant relationship to WTP. Migraine sufferers who already paid for prescription treatment were more likely to be willing to pay than those who did not use prescription treatment [87].

Four studies included a scope variable to test whether the WTP was sensitive to the size of health gains [72, 79, 86, 89]. Two studies showed a significant association with WTP. A complete abolition of the risk of restenosis was associated with higher WTP of patients participating in a WTP study alongside a clinical trial [86]. Higher reduction level in anginal attacks significantly increased the WTP in patients with angina pectoris [89].

Miscellaneous factors were included in 11 studies. As they referred mainly to methodological issues such as, e.g. the influence of starting bids or visual aids, they were not beyond the scope of this review.

The explanations offered by the authors to clarify the meaning of the identified predictors of WTP varied widely. Not all authors gave an interpretation of

their findings. Appendix 2 shows a summary of explanations given for significant predictors.

3.4 Discussion

The results of this review indicate that there were a wide range of factors predicting the WTP for changes in pain and related symptoms beyond income. Several key issues emerged.

Firstly, the evidence with regard to different predictors varied widely. Some predictors, such as age and gender, were frequently examined and were included in 17 and 15 studies, respectively. All of the other predictors were only analyzed in a few studies, even pain was examined in only six studies.

No study simultaneously examined all of the predictors identified by this review, so their relative importance is as yet not understood.

Secondly, based on the ICF as an integrative model of functioning, disability and health, the predictors could be meaningfully structured. They included factors covering almost all of the ICF components. Predictors related to functional impairments contained aspects such as depression, anxiety, pain, and angina pectoris attacks. Limitations in activities and participation were represented by difficulties in walking and daily living. Contextual factors were addressed by environmental factors such as care in a public clinic. Personal characteristics, which mainly included socio-demographic features comprised additional aspects including

coping skills, experiences with the disease or treatment, and knowledge about the disease.

The predictors encompassed ICF components as well as factors related to general health, such as the current perception of health, and identified included factors referring to the health condition, such as the severity of the underlying disease. In addition, the predictors represented aspects of satisfaction with healthcare that included recommending a treatment or an opinion about the healthcare system. Finally, the predictors covered health expenditures such as copayments and included so-called scope variables, such as risk reductions.

The broad range of predictors identified was not surprising, as it mirrored the various health conditions of included studies eliciting the WTP for changes in pain and related symptoms. Moreover, it showed that there were different aspects of a disease incorporated in the valuation of outcomes such as pain and related symptoms.

A third main finding was that the current understanding of the identified predictors of WTP varied widely. Obviously, factors representing respondents' current functioning as well as their general health or severity of illness were assumed to influence the WTP in the context of valuing health outcomes. In many studies, pain and related symptoms predicted the WTP. In other words, patients experiencing, e.g. more limitations in daily activities were willing to pay higher amounts for a cure of their arthritis [93]. In that study, WTP captured one of the most relevant aspects of this condition, since together with pain, restrictions in activities and participation may be most relevant to patients with rheumatoid arthritis [96]. Furthermore, the

Thompson study, as well as other studies, showed that the WTP increased as functioning decreased. For example, patients with more weekly angina attacks had a higher WTP for a defined reduction in angina pectoris attacks [89]. Such relationships can also be observed for many other outcomes, such as the relief or cure of e.g. asthma, psoriasis or menopausal symptoms [27, 97, 98]. The hypothesis proven in these studies was that patients with more severe symptoms were willing to pay more for a hypothetical cure than those with less severe symptoms. Significant associations between severity of illness and WTP were then considered to contribute to the construct validity of the WTP method.

As this review showed, not all studies found the relations between WTP and functioning to be in the expected direction. For example, in a study about the WTP for a perfect migraine drug, moderate migraine disability (MIDAS Grade III) and not severe disability (MIDAS Grade IV) was what affected the WTP significantly [87]. The authors argued that MIDAS Grade IV sufferers are more likely to be depressed and may have given up on the possibility of excellent relief. Interesting to note is that, in another study, more severe depressive symptoms evoked a decline in the WTP for a cure of rheumatoid arthritis [93], but no explanation was offered by the author. Psychological factors may possibly moderate the relationship between functioning status and WTP. Further research in this area will be useful and is strongly recommended.

Some of the predictors of WTP as identified in this review, may serve rather as indicators of recognized predictors of WTP. For example, the influence of age was considered as a proxy for health in two reviewed studies [81, 91]. The indicator function of age was also suggested in meta-analyses of WTP estimates of morbidity

studies [99, 100]. However, the influence of such proxy variables may disappear if the underlying predictor can be adequately controlled for. This may question their role as independent predictors, except in cases where data on a known predictor are not available. Satisfaction was considered in some studies as a proxy for the treatment effect. As pointed out in one reviewed study, patients' overall opinion of treatment was suggested to affect the WTP because it reflected the clinical benefits derived by the patients [92]. This is in line with literature regarding satisfaction as a measure of treatment success which may affect the WTP [101]. Patient satisfaction is a recipient's rating of the results of his or her treatment experience and of salient aspects of process [102].

Some predictors identified from the set of personal factors may be explained for very different reasons; experiences with the treatment, for example, were explained in one reviewed study to lower the WTP for reduced risk of restenosis. It was suggested that this was because patients with a previous bypass surgery were less concerned about the possibility of a second operation [86]. Manifold reasons as to why familiarity with a disease or treatment influenced the WTP were pointed out in another study not included in this review. Individuals' familiarity with lymphatic filariasis was suggested to lower the WTP for prevention and treatment due to a) coping and increased acceptance, b) perceived financial constraints, c) lack of belief in successful treatment, and d) resignation [103].

Ethnic origin was also a predictor discussed with regard to possible reasons for its influence on WTP. In this study showing a predictive value for ethnicity, the authors speculated that ethnic groups may place different values on difficulty in walking and mobility [80]. Moreover, lack of awareness or experience with benefits of

treatment or lack of trust in the medical system were suggested to lower the value of health improvements in osteoarthritis for African Americans. Literature from previous research provided further knowledge concerning the influence of ethnicity on WTP. Beliefs about fatalism were suggested as a possible explanation for the lower WTP of Filipino and Chinese women for mammography [104]. Variables such as ethnicity or experiences are easy to elicit, but they provide no information on the etiology of differences. Therefore, their inclusion in studies may be a first step to look at differential variables of WTP.

Several limitations of this review should be noted. The search was confined to articles published in English which may have limited the number of included studies and the range of potential predictors. However, because there was only one not-English written study with an English abstract providing predictors of WTP similar to the included studies [105], the exclusion is believed not to impact the conclusions to be drawn from this review. Despite using a thorough search strategy, there may be some literature on predictors of WTP that were not identified for this review. Specifically, the grey literature (unpublished documents) on this topic was not examined, focusing instead on publications that had been through the peer-review process. Furthermore, a lack of specific MeSH (Medical Subject Headings) terms for this topic made it difficult to formulate a search strategy that was both specific and sensitive in retrieving articles for this review. Direct comparison of results was limited by the heterogeneity of the studies. The studies included differed regarding study populations, research settings, medical problems, health care systems, sample sizes, estimation methods, and inclusion of other variables. Therefore, it is not possible to justify a recommendation on a single predictor because the quantity and consistency of results cannot be adequately assessed at this time.

In summary, a broad range of predictors was revealed by the available evidence. All identified predictors affected the WTP beyond income and represented relevant personal characteristics such as an individual's current functioning or socio-demographic background. The findings of the various predictors suggest that answers to hypothetical WTP questions are reasonable and meaningful, and not simply random numbers. These general findings are a synthesis drawing upon the specific findings of many different WTP studies. More important in practice is to look at the relationships in each particular study and to find out whether the predictors can explain much of the variation in respondent's WTP. Only in cases where a sufficient proportion of variation in WTP is explained the researcher understands the factors that drive the respondent's WTP. Furthermore, in contrast to most studies that have examined the relevance of a few factors at a time, the simultaneous examination of all the identified predictors enables the determination of their relative importance which is as yet not understood. Therefore, predictors of WTP for changes in pain and related symptoms should preferably be investigated in future studies from a comprehensive perspective.

Table 1: Summary of reviewed WTP studies controlled for income (N=21)

Study	Year of publication	Country	Objective	Population description	Sample size	Survey method	Elicitation format	Statistical analysis
Allen et al. [76]	2007	USA	WTP for a hypothetical muscle relaxant that completely prevents postoperative myalgia, reduces its incidence to 3%, and not associated with myalgia, bradycardia, malignant hyperthermia	adult patients scheduled for elective surgery during general anesthesia	88	computerized self-administered questionnaire	bidding game	analysis of variance
Atchison et al. [77]	2007	USA	WTP to have jaw fracture fixed with instant return to normal function and with no chance of permanent numbness or scar	patients receiving treatment for either a mandibular fracture or third molar removal	203	face-to-face interview	payment card	linear regression
Bala et al. [78]	1998	USA	WTP to move from one pain-duration profile to another	persons in the 65- to 70-year - old age group in Florida	114	computer based face-to-face interview	dichotomous choice	logistic regression
Boonen et al. [79]	2005	Austria, the Netherlands	WTP for treatment in a spa resort of patients with ankylosing spondylitis	2 intervention groups (n=80) with a 3 week inpatient program and a control group (n=40) with NSAR and physical exercise	120	self-administered questionnaire	payment card	ordinal logistic regression
Byrne et al. [80]	2004	USA	WTP for improvement in severe and mild Osteoarthritis	race/ethnic stratified sample of adult residences of Harris County, Texas	193	face-to-face interview	open-ended	linear regression
Chestnut et al. [72]	1996	USA	WTP to avoid additional angina symptoms	men with a history of chest pain and a physician's diagnosis of angina pectoris	35	face-to-face interview	open-ended; dichotomous choice	linear regression
Cross et al. [81]	2000	Australia	WTP for joint replacement	patients 2-3yr after their total hip replacement (n=109) or total knee replacement (n=129)	238	self-administered questionnaire	payment card	logistic regression
Donaldson et al. [82]	1997	Scotland	WTP for laparoscopic treatment	patients on the waiting list for cholecystectomy at Aberdeen Royal Hospitals NHS Trust	117	self-administered questionnaire	payment card	linear regression

Study	Year of publication	Country	Objective	Population description	Sample size	Survey method	Elicitation format	Statistical analysis
Dranitsaris [83]	1999	Canada	WTP for bisphosphonate (pamidronate)	Canadian tax-payers with permanent residence status in Ontario	100	face-to-face interview	payment card	linear regression
Ethgen et al. [84]	2003	Belgium	WTP for a treatment yielding relief or a total cure of osteoarthritis	patients with hip and knee osteoarthritis attending the outpatient physical rehabilitation and rheumatology clinic of the University Hospital of Liege	128	face-to-face interview	dichotomous choice, bidding game	logistic, linear regression
Fautrel et al. [85]	2005	Canada	WTP for a 100% effective cure of rheumatoid arthritis	convenience sample of rheumatoid arthritis patients from 5 participating rheumatologists affiliated with the McGill University Health Center Division of Rheumatology	121	telephone interview	payment card	ordered logit regression
Greenberg et al. [86]	2004	USA	WTP for avoiding a repeat revascularization procedure	patients who underwent percutaneous coronary interventions as part of two multicenter randomized trials	1642	self-administered questionnaire	dichotomous choice	logistic regression
Hamelsky et al. [87]	2005	USA	WTP for acute medication for severe headache attacks	population-based database of headache sufferers	201	self-administered questionnaire	payment card	logistic regression
Jacobs et al. [88]	2002	USA	WTP for preventing hepatitis A symptoms	sample of American adults from a national mailing list from motor vehicle and real estate records	178	self-administered questionnaire	payment card	general linear regression
Kartman et al. [89]	1996	Sweden	WTP for a more effective drug reducing weekly anginal attacks by 50%	angina pectoris patients	341	telephone interview	dichotomous choice, bidding game	logistic, linear regression
King et al. [90]	2004	USA	WTP for a cure for all health problems and symptoms with a single payment	patients with cervical spondylotic myelopathy at a Veterans Affairs neurosurgery clinic	79	face-to-face interview	bidding game	linear regression

Study	Year of publication	Country	Objective	Population description	Sample size	Survey method	Elicitation format	Statistical analysis
Narbro et al. [91]	2000	Sweden	WTP for a treatment that relieves overweight-related problems	obese men (1479) and women (2070) of the Swedish Obese Subjects registry study	3549	self-administered questionnaire	open-ended	linear regression
Ritchie [92]	1996	UK	WTP for preferred treatment of soft-tissue rheumatic conditions of the shoulder or elbow	outpatients	126	self-administered questionnaire	payment card	linear regression
Thompson [93]	1986	USA, Canada	WTP for a complete cure of arthritis	patients with rheumatoid arthritis	247	face-to-face interview	open-ended	linear regression
van den Bosch et al. [94]	2006	The Netherlands	WTP for perfect antiemetics or analgesics	mix of adult inpatients and outpatients	808	self-administered questionnaire	bidding game	ordinal logistic regression, linear regression
Werner et al. [95]	2002	Israel	WTP for drug treatment for osteoporosis, reducing the risk of hip fracture by 50%	postmenopausal women recruited from an outpatient clinic	109	telephone interview	open-ended	linear regression

WTP = willingness to pay

Table 2: Predictors of WTP classified by the ICF components, 1st or 2nd level ICF categories

	First author																	prediction				
	Allen	Atchison	Bala	Boonen	Byrne	Chestnut	Cross	Donaldson	Dranitsaris	Ethgen	Fautrel	Greenberg	Hameisky	Jacobs	Kartman	King	Narbro		Ritchie	Thompson	van Bosch	Werner
ICF COMPONENT																						
ICF Category																						
2nd level ICF category																						
BODY STRUCTURES																						
s7 Structures related to movement																						
s770 Additional musculoskeletal structures related to movement																					-	0/1
BODY FUNCTIONS																						
Body functions																						
b1 Mental functions																						
b114 Orientation functions				-																		1/4
b126 Temperament and personality functions																					-	0/1
b130 Energy and drive functions												-										0/1
b152 Emotional functions																			+/-		+	2/2
b2 Sensory functions and pain																						
b280 Sensation of pain																						2/6
b4 Functions of the cardiovascular, haematological, immunological and respiratory systems																						
b410 Heart functions																						2/2
b7 Neuromusculoskeletal and movement-related functions																						
b710 Mobility of joint functions																						0/2
ACTIVITIES & PARTICIPATION																						
Activities & participation																						
d4 Mobility																						
d450 Walking																						2/5
																						1/1

	First author																		prediction			
	Allen	Atchison	Bala	Boonen	Byrne	Chestnut	Cross	Donaldson	Dranitsaris	Ethgen	Fautrel	Greenberg	Hamelsky	Jacobs	Kartman	King	Narbro	Ritchie		Thompson	van Bosch	Werner
ICF COMPONENT																						
ICF Category																						
2nd level ICF category																						
ENVIRONMENTAL FACTORS																						
e1 Products and technology																						
e110 Products or substances for personal consumption																						
e165 Assets ¹⁾ (income)	+	+	-	-	+	+	+	+	-	+	-	+	-	+	+	+	+	-	+	+	-	0/1 14/21
e5 Services, systems and policies																						
e570 Social security services, systems and policies																						
e580 Health services, systems and policies																						
PERSONAL FACTORS²⁾																						
Attitudes & Beliefs																						
Coping																						
Experiences																						
Knowledge																						
Other health conditions																						
Physical characteristics																						
Traits																						
Sociodemographic																						
Age	-	-	-	-	+		+		-	-	+		-	-	+		+	-	+	-	-	6/17
Education																						
Employment																						
Gender																						
Household size																						
Insurance																						
Marital status																						
Race																						
Religion																						
Social class																						

	First author																						
	Allen	Atchison	Bala	Boonen	Byrne	Chestnut	Cross	Donaldson	Dranitsaris	Ethgen	Fautrel	Greenberg	Hamelsky	Jacobs	Kartman	King	Narbro	Ritchie	Thompson	van Bosch	Werner	prediction	
ICF COMPONENT																							
ICF Category																							
2nd level ICF category																							

+ Predictor significant in multivariate analysis ($p \leq 0.05$)

- Predictor not significant in multivariate analysis

+/- If more than one predictor refers to the same ICF component or ICF 1st or 2nd level category, but one of these predictors is significant and the other not e.g. knowledge of osteoporosis and knowledge of regulations (Werner)

Absence of symbol indicates predictor was not assessed

¹⁾ All assets (income) such as money, property, and other valuables that an individual owns

²⁾ Personal factors are not currently classified in the ICF

Table 3: Predictors of WTP not linked to the ICF

	First author																			prediction			
	Allen	Atchison	Bala	Boonen	Byrne	Chestnut	Cross	Donaldson	Dranitsaris	Ethgen	Fautrel	Greenberg	Hamelsky	Jacobs	Kartman	King	Narbro	Ritchie	Thompson		van Bosch	Werner	
Health condition					-					-					+		+/-		-		-	2/6	
General health								-									-	-	+/-				1/4
Expenditures (health related)						-							+							-			1/3
Satisfaction with healthcare services								+/-			+								+				3/3
Scope				-		-							+			+							2/4
Miscellaneous			+	-		+				+	-	+		-	+					-	+	+	7/11

+ Predictor significant in multivariate analysis ($p \leq 0.05$)

- Predictor not significant in multivariate analysis

+/- If more than one predictor refers to the same predictor domain, but one of these predictors is significant and the other not e.g. General health / RAND current health and RAND prior health (Thompson)

Absence of symbol indicates predictor was not assessed

4 An exploratory analysis of predictor candidates of willingness to pay for small to moderate health effects of rehabilitation interventions

4.1 Objective

The objective of this part of the doctoral thesis was to explore factors beyond income and scope that predict the WTP based on contingent valuation for expected and perceived small to moderate health effects of rehabilitation interventions. The specific aims were (1) to examine a comprehensive set of putative predictors of WTP, and (2) to investigate whether they vary across health conditions.

4.2 Methods

4.2.1 Study population

Data were collected from consecutive patients with five different health conditions. Patients were recruited from two centers: patients with musculoskeletal disorders (osteoarthritis, osteoporosis, back pain) undergoing outpatient or day clinic rehabilitation from the Department of Physical Medicine and Rehabilitation of the University of Munich Hospital, and patients with psychosomatic disorders (depression, pain disorder) from inpatient rehabilitation at the Medical Psychosomatic Clinic Roseneck in Prien / Chiemsee. This longitudinal study was performed between January 2003 and March 2005. For a detailed description of the study design, see Brach M, Gerstner DG, Stucki G (Usefulness of the WTP method to evaluate small to moderate clinical effects in rehabilitation, to be provided for publication in 2010).

The majority of the study population was female (68% of osteoarthritis and back pain, 85% of osteoporosis patients, 71% of pain disorder and 64% of depression patients) with a mean age of approximately 67 years for both osteoarthritis and osteoporosis, and 53 years for back pain patients. Psychosomatic patients were younger and, on average, 47 years old (see Table 1). Of the 539 participants at the beginning of rehabilitation, 92 were not eligible at the end of rehabilitation because they refused to take part or dropped out due to medical reasons or a premature termination of therapy. A withdrawal analysis for each health condition showed no significant difference in baseline characteristics between participants and non participants.

Patients were included if they were at least 18 years old, had sufficient knowledge of the German language, had the cognitive ability to understand the purpose of the study, and signed an informed consent form. The study protocol and informed consent forms were approved by the Ethics Committee of the Ludwig-Maximilians-University of Munich.

4.2.2 Data

All patients completed structured interviews conducted by trained interviewers. The survey elicited the WTP for defined health effects and personal available income.

Additionally, patients completed several self-administered questionnaires measuring functioning and personal factors.

4.2.3 Willingness to pay measurement

WTP was elicited by the contingent valuation method, in which respondents state monetary values for health benefits “contingent” on a given hypothetical scenario. Using the payment card format in this study, patients were asked what maximum monthly (€) amount they would be willing to pay for both an expected and a perceived health effect. The expected health effect was defined as the difference between the baseline pain intensity and the pain intensity patients expected after rehabilitation. Pain intensity was measured on a VAS ranging from 0 (no pain) to 10 (unbearable pain). This scenario was presented to the patients before starting treatment. At the end of rehabilitation, the patients were introduced to a second scenario, the perceived health effect, which was the difference between their baseline and current pain intensity. If the patient had not experienced pain during the last seven days, then the general health status was measured on a VAS (0/10; 0 = worst health; 10 = best health). The detailed interview has been described by Brach M, Gerstner DG, Stucki G (Usefulness of the WTP method to evaluate small to moderate clinical effects in rehabilitation, to be provided for publication in 2010).

4.2.4 Predictors

Patients completed a series of self-report measures at the beginning of the first treatment session. Socio-demographic variables were recorded, including gender, age, marital status, education level, job qualification, employment status, health insurance and household income. The assessment of functioning status included the eight multi-item scales of the Short Form-36 (SF-36) [106, 107]. The

score of the Self-administered Comorbidity Questionnaire (SCQ) was used to evaluate the burden of coexisting conditions [108].

Specific aspects of the respective health condition were measured in osteoarthritic patients using the three Western Ontario and McMaster Universities (WOMAC) scales [109, 110], in osteoporotic patients by the Stanford Health Assessment Questionnaire (HAQ) functional disability index [111, 112], and in back pain patients by the two subscales of the North American Spine Society Lumbar Spine Outcome Assessment instrument (NASS) [113, 114]. To evaluate the severity of illness of depressive patients the Allgemeine Depressionsskala (ADS) [115, 116] was used. The two subscales of the Pain Perception Scale (Schmerzempfindungsskala - SES) [117, 118] were used to measure the burden of disease of patients suffering from pain disorder, and were additionally applied to all patients suffering from musculoskeletal conditions.

All patients completed instruments measuring their beliefs about control over one's health and strategies to cope with stress. The German version of the Health Locus of Control (Fragebogen zur Erhebung von Kontrollüberzeugungen zu Krankheit und Gesundheit - KKG), that includes the three scales internality, externality, and fatalism [119, 120], was used to elucidate health control beliefs. Skills that characterize proactive coping behaviors to deal with stressful events were assessed using the two subscales of preventive coping and strategic planning of the Proactive Coping Inventory (PCI) [121]. Except for depressive patients, all other participants' cognitive and behavioral strategies to cope with pain were evaluated by the revised Coping Strategies Questionnaire (CSQ_R) [122, 123]. Two single items

that assess patients` subjective ability to control and decrease their pain were also included.

The contents of the scales and meaning of the scores are described in appendix 3.

The International Classification of Functioning, Disability and Health (ICF) of the World Health Organization [46] was used as a reference framework to classify all measures. The ICF covers relevant domains encountered in patients in need of rehabilitation [73]. Table 2 shows the measures and their distribution over the ICF components Body Functions, Activities and Participation, Environmental and Personal Factors. An additional domain, General Health, encompassed measures which addressed general health aspects not covered by the ICF components.

As a quality control measure all submitted questionnaires were reviewed thoroughly for completeness, accuracy, and internal consistency of data. Validation of 10% of data entry was done by entering the data again.

4.2.5 Statistical analysis

The WTP for the expected, and the perceived health effects were used as response variables to explore the above mentioned set of potential predictors. Unless noted otherwise, all steps of analysis were done for both the WTP for expected and perceived health effects.

In order to handle a large number of potential predictors two statistical approaches were used which selected predictors in different ways: the content-oriented stepwise-built generalized linear regression modelling (GLM) and the data-driven least absolute shrinkage and selection operator (LASSO).

Content-oriented stepwise-built generalized linear regression modelling (GLM)

Predictor candidates emerged from regression models which were constructed in three steps: (1) testing significant Spearman correlations with WTP ($p < 0.05$), (2) screening for significant predictors ($p < 0.05$) within separate GLMs for the three ICF components Body Functions, Activities and Participation, and Personal Factors, and (3) building a final model with significant predictors derived from step 2 across all ICF components. In the presence of bivariate collinearity among predictors ($\rho > 0.5$), alternative models exchanging the correlated predictors were built. All GLMs were further controlled for disposable income and scope, a predictor candidate was considered as significant if $p < 0.05$. All possible final models were checked for multi-collinearity by the variance inflation factor (VIF).

Due to the peculiarity of the distribution of WTP, having a heavy right tail, a generalized linear regression model assuming a gamma-distribution of WTP and a log-link between WTP and the predictors was used. As a measure for the variation in the response explained by the model, Nagelkerke's R-square, an analogue to R-square in linear regression, was employed.

Least absolute shrinkage and selection operator (LASSO)

The LASSO is an extension of the linear model [124]. By imposing a restriction on the regression parameters the LASSO algorithm inherits an implicit

variable selection property, i.e., it can shrink some of the regression coefficients to zero. The LASSO thus does not require a pre-specified set of predictors but rather selects only the most influential predictors that contribute to the understanding of WTP.

The results depend on the degree of shrinkage λ . Using a cross-validation procedure we identify the optimal value for λ and thus minimize the corresponding mean squared error.

The LASSO cannot deal well with missing values in the predictors so an imputation technique was used to replace missing values by a set of estimates [125]. Each missing value was thereby replaced by a total of 10 attempts to reconstruct the original value. Consequently 10 different imputed data sets that reflected the randomness of the imputation procedure and also 10 sets of predictors from the LASSO regression, one for each data set were obtained. A predictor was considered as relevant for the understanding of WTP if its corresponding coefficient deviated from 0 in at least 7 of the 10 imputed data sets.

Descriptive analyses were performed utilizing the SPSS, version 15.0 (SPSS, Chicago, IL). For generalized linear regression analyses the SAS version 9.1 (SAS Institute Inc., Cary, NC) was used. The LASSO regression was provided by the `grplasso` package for R version 2.8.1 (R foundation for statistical computing, Vienna, Austria).

4.3 Results

The two statistical methods returned two non-overlapping but, by themselves, reasonable sets of predictor candidates for the understanding of WTP beyond income and scope. Some of the predictors were relevant for WTP for both the perceived and expected health effects.

4.3.1 LASSO

By following the frequency criterion, i.e. the appearance of a predictor in at least 7 out of 10 imputed models, additional factors predicting WTP were identified. Table 3 shows the median multiplicative effect of a predictor on the WTP over all imputed samples, denoted by $\exp\{\beta_{\text{med}}\}$. WTP was modelled on the log scale. Thus, increasing a predictor by one unit had a median effect on WTP equal to the multiplicative factor $\exp\{\beta_{\text{med}}\}$. For example, if SF-36 role-physical increased by one unit for back pain patients, on average a decrease of 0.46% in the WTP for expected effects and of 0.35% for perceived effects was observed. Note that high or low scores do not have the same meaning across different scales: for the SF-36 scales a higher score represents better health. The transformed coefficients, $\exp\{\beta_{25}\}$ and $\exp\{\beta_{75}\}$ corresponded to the 25% and 75% percentiles, respectively, and thus served as a measure of uncertainty. The predictors stemmed from the ICF components Body Functions and Activities and Participation. Especially noteworthy was the frequent appearance of subscales of the SF-36 such as bodily pain, physical functioning, role-emotional and role-physical. Predictors were found for either the WTP for expected or perceived effects; the same predictor for both WTP measures emerged only for depression and back pain patients.

4.3.2 Stepwise-built GLM

The ICF based GLM resulted in the identification of a different set of predictors of WTP (Table 4). Similar to the LASSO, $\exp\{\beta\}$ represented the mean multiplicative effect of an increase in a predictor on the WTP. Worthy of note is that most parameter estimates were much larger than for the LASSO regression. This simply followed from the underlying scaling of the parameters (e.g. school education as binary, SF-36 on a 0-100 scale). The predictors represented all three ICF components Body Functions, Activities and Participation and Personal Factors. School education was the dominant factor, predicting the WTP a total of four times. School education and HAQ were the factors predicting the WTP for both expected and perceived effects. The same predictor for both WTP measures was observed only for osteoporosis and osteoarthritis. The R^2_{final} indicated the improvement as compared to a baseline model containing only the intercept, disposable income, and scope. The collinearity statistics revealed that the VIFs of the final models ranged from 1.002 to 1.330, and thus multi-collinearity might not have been a serious problem in the GLM regression models [126].

4.4 Discussion

This study demonstrated that there were additional predictors beyond income and scope which help to explain what drives the WTP for small to moderate health effects in the context of rehabilitation interventions. From a comprehensive set of potential predictors it was possible to identify predictors which were relevant for WTP. This was either due to their frequent appearance in LASSO or due to their significance and ability to explain variance in GLM. The explained additional variance

of the significant predictors varied between 2% and 9% and enhanced the explanatory power of base models containing only income and scope up to a total explained variance ranging from 23% to 42%. The relevant predictors in the LASSO models appeared in at least 7 out of 10 imputed datasets.

Using the ICF as a reference predictors were included within a sound conceptual framework. The identified predictor candidates included the ICF components Body Functions, Activities and Participation, as well as Personal Factors. This wide range of relevant predictors was not surprising, since it mirrored the complexity of rehabilitation medicine, which focused on limitations of functioning and disability associated with health conditions and with the complex interaction with personal factors and the environment [47]. The spectrum of identified predictors was also in line with the systematic review of predictors of WTP for changes in pain and related symptoms reported in the first part of this doctoral thesis, where reference is made to different health conditions such as, e.g. angina pectoris, myalgia or rheumatic diseases. The fact that in this study virtually the same domains were found to be relevant as were elucidated in studies referring to different areas supports the strategy of including predictors from a content-oriented perspective. Based on the ICF the concepts of important predictors of WTP are thus translated into a common language for health and disability, irrespective of the underlying health condition [46, 127]. This improves the understanding of the drivers of WTP and their comparison across different health conditions and studies.

This study showed that different predictors of WTP seem to matter for different health conditions. While some predictors were predictive in multiple health conditions, including school education or limitations in daily activities due to physical

problems as measured by the SF-36 role-physical, others affected the WTP in specific conditions. For example, psychological aspects such as coping appeared to influence the WTP of back pain patients, and depressive symptomatology played a role in patients with depression. From our current understanding of the key aspects of the examined health conditions the identified predictors showed face validity. Research has suggested that psychological factors such as coping skills may influence susceptibility to chronic pain and pain control [128, 129] and should be targeted in the primary care management of back pain [130] and in rehabilitation to improve outcomes [131]. Depressive symptoms, including e.g. feelings of guilt and worthlessness or loss of appetite, are usually measured to assess whether a person is depressed or not [116].

How can these results be used in future studies? Based on the findings, it could be recommended that future studies examining the WTP for small to moderate health effects in the area of functioning, disability and health should take into consideration the inclusion of predictors addressing aspects of functioning and personal background relevant to specific conditions. For this purpose, future studies may relate to the ICF Core Sets which have been developed for specific chronic health conditions [132]. Furthermore, when studying the impact of functioning on WTP the use of a generic instrument such as the SF-36 or the ICF-based WHODAS-II (WHO Disability Assessment Schedule II) [133] can be recommended. As shown in this study, many aspects of functioning were well reflected by the SF-36 subscales. Except for very specific aspects, including depression symptoms or pain sensations, functioning as represented by the SF-36 was able to predict WTP. Disease specific scales measuring similar concepts provided only marginal value. For example, the HAQ predicted WTP efficiently well in osteoporosis patients and additionally

explained a 6% variance in WTP. The SF-36 subscale physical functioning, which was highly correlated with the HAQ ($\rho = -0.726$), also emerged as a predictor of WTP in an alternative model specification and provided a comparable, explained variance of 4%. Regardless of whether a disease-specific or a generic measure of functioning is used as a potential predictor of WTP, the administration of a generic measure has the advantage of allowing comparisons across different conditions and studies [134].

With the LASSO and the stepwise-built GLM we compared two statistical procedures that covered two largely varying settings. Whilst for the LASSO we followed a completely data-driven approach, for the GLM we ensured the absence of collinearity and built the model from a content-oriented perspective. As the results showed, the two different methods yielded two almost non-overlapping sets of predictors which, by themselves, seemed to be reasonable. Hence, it is important to qualify the observed lack of overlap. The different findings may be partly due to divergent underlying assumptions of the applied statistical techniques. The LASSO cannot deal with missing values (requiring imputation) and assumes a normal distribution with log-link, whereas the GLM assumes a log-link and a Gamma distribution for the WTP. Moreover, as a consequence of the high degree of multicollinearity in the data, predictors might, to a certain degree, be interchangeable. The sensitivity of WTP results to the researcher's choice of estimation model has already been described elsewhere [135, 136]. Therefore we tentatively suggest that model specification might also be responsible for the observed discrepancies in the two identified sets of predictors of WTP. More research is warranted to investigate the effect of assumptions underlying statistical methods on predictor selection.

Clearly, given the study's limitations, the findings from this exploratory study should be viewed as preliminary. One major design limitation was sample size which was comparatively small in relation to the number of potential predictors, this then limited the ability to detect significant results. While patients were recruited from two different facilities, they nonetheless represented a geographically homogenous group in each health condition, as patients with the same health condition stemmed solely from one facility. This caused concerns related to the issue of generalizability.

Furthermore, the comprehensive set of questionnaires, resulting in a large number of items to fill in, may have evoked a response burden on the patients and may, to some extent, explain missing values [137, 138]. Though the missing items were observed to be mostly below 10%, in some cases the missing rate was higher, albeit always less than 20%, which is assumed not to have a major impact on study findings [139].

Lastly, two WTP measures were used, the WTP for expected and for perceived effects, in order to find a larger quantity of potential predictor candidates. An *a priori* hypothesis was not stated as to whether or not predictors should be the same for both scenarios. As the results showed, there were some similarities, but more frequently, different predictors were found. It could not, therefore, be precluded that using either the expected or perceived WTP had an effect on the resulting importance of predictors.

In conclusion, beyond income and scope, there were additional factors predicting the WTP for small to moderate health effects as encountered in rehabilitation interventions. They included the ICF components Body Functions,

Activities and Participation, and Personal Factors. They were relevant and differed for different conditions. Except for specific aspects of a condition, predictors representing functioning might be well captured by a generic health such as the SF-36, which provides data comparable across conditions and studies. The influence of school education should be checked. Just how sensitive empirical results were became obvious with respect to changes in the underlying model structure.

4. An exploratory analysis of predictor candidates for willingness to pay

Table 1: Demographics of the study population (n = 539) at the beginning of rehabilitation

Characteristics	Osteoarthritis (n = 97)	Osteoporosis (n = 98)	Back pain (n = 143)	Pain disorder (n = 96)	Depression (n = 105)
Age mean ± SD years	66.6 (10.5)	66.9 (7.6)	52.9 (14.2)	46.8 (10.5)	46.6 (11.5)
Gender					
Male	31 (32.0)	15 (15.3)	46 (32.2)	28 (29.2)	38 (36.2)
Female	66 (68.0)	83 (84.7)	97 (67.8)	68 (70.8)	67 (63.8)
Marital status					
Single	9 (9.3)	22 (22.4)	32 (22.4)	18 (18.8)	24 (22.9)
Married	51 (52.6)	48 (49.2)	84 (58.7)	55 (57.3)	60 (57.1)
Divorced/Separated	10 (10.3)	16 (16.3)	8 (5.6)	18 (18.8)	19 (18.1)
Widowed	24 (24.7)	8 (8.2)	5 (3.5)	5 (5.2)	1 (1.0)
School education					
Elementary school	40 (42.6)	31 (33.0)	38 (29.9)	37 (38.9)	34 (33.0)
Junior high school	27 (28.7)	38 (40.4)	31 (24.4)	19 (20.0)	26 (25.2)
Senior high school	22 (23.4)	21 (22.3)	54 (42.5)	38 (40.0)	39 (37.9)
Other	5 (5.3)	4 (4.3)	4 (3.1)	1 (1.1)	4 (3.9)
Job qualification					
Qualified job	38 (41.3)	36 (40.0)	36 (28.3)	25 (26.3)	30 (29.1)
Technical college graduate	21 (22.8)	28 (31.1)	43 (33.9)	38 (40.0)	41 (39.8)
University graduate	15 (16.3)	6 (6.7)	27 (21.3)	20 (21.1)	18 (17.5)
Other	18 (19.6)	20 (22.2)	21 (16.5)	12 (12.6)	14 (13.6)
Employment status					
Employed	18 (19.1)	10 (10.6)	77 (60.6)	61 (64.2)	68 (64.8)
Unemployed	12 (12.8)	13 (13.8)	13 (10.2)	14 (14.7)	21 (20.0)
Retired	62 (66.0)	71 (75.5)	35 (27.6)	16 (16.8)	14 (13.3)
Other	2 (2.1)	0	2 (1.6)	4 (4.2)	2 (1.9)
Insurance status					
Statutory	72 (76.6)	66 (71.0)	85 (67.5)	59 (62.8)	47 (44.8)
Statutory & additionally privat	17 (18.1)	23 (24.7)	24 (19.0)	11 (11.7)	15 (14.3)
Privat	5 (5.3)	4 (4.3)	17 (13.5)	24 (25.0)	43 (41.0)
Net household income €					
≤ 500	1 (1.2)	2 (2.2)	0	3 (3.5)	2 (2.0)
500 to 1000	8 (9.3)	10 (11.1)	6 (4.9)	4 (4.7)	7 (7.1)
1000 to 1500	25 (29.1)	27 (30.0)	16 (13.1)	13 (15.1)	11 (11.2)
1500 to 2000	16 (18.6)	21 (23.3)	22 (18.0)	13 (15.1)	15 (15.3)
2000 to 2500	12 (14.0)	11 (12.2)	25 (20.5)	13 (15.1)	13 (13.3)
2500 to 3000	9 (10.5)	12 (13.3)	24 (19.7)	8 (9.3)	8 (8.2)
3000 to 3500	6 (7.0)	4 (4.4)	10 (8.2)	10 (11.6)	9 (9.2)
≥ 3500	9 (10.5)	3 (3.3)	19 (15.6)	22 (25.6)	33 (33.7)
Monthly disposable income € mean ± SD years	518.1 (631.0)	405.4 (392.9)	487.7 (492.9)	407.1 (435.1)	439.9 (433.7)

Except where indicated otherwise, values are the number (%) of patients

4. An exploratory analysis of predictor candidates of willingness to pay

Table 2: Measures and their distribution over the ICF components

Measures	Subscales/item	ICF Component				Not covered by the ICF General Health
		<i>Body Functions</i>	<i>Activities & Participation</i>	<i>Environmental Factors</i>	<i>Personal Factors*</i>	
ADS	General Depression Scale	+				
HAQ	Standard Disability Index		+			
NASS	Pain and disability scale Neurogenic scale	+	+			
WOMAC	Pain	+				
	Stiffness	+				
	Function		+			
SF-36	Physical functioning		+			
	Role limitation due to physical problems		+			
	Bodily pain	+				
	General health perception					+
	Vitality	+				
	Social functioning		+			
	Role limitation due to emotional problems		+			
	Mental health	+				
SES	Health change					+
	Affective	+				
HLC	Sensory	+				
	Internality				+	
PCI	Externality				+	
	Chance				+	
	Preventive coping				+	
CSQ_R	Strategic planning				+	
	Distraction				+	
	Distancing from pain				+	
	Coping self statements				+	
	Ignoring pain				+	
	Praying				+	
	Catastrophizing				+	
	Pain control				+	
Pain decrease				+		
SCQ	Score comorbidity				+	
Socio-economic questionnaire	Age				+	
	Gender				+	
	School education				+	
	Job qualification				+	
	Employment status				+	
	Health insurance				+	
	Marital status				+	
Income (Interview)	Household income			+		
	Disposable income			+		

* Personal Factors not currently classified in the ICF

+ Measure/item is linked to the ICF

ADS = Allgemeine Depressionsskala (General Depression Scale); HAQ = Health Assessment Questionnaire; NASS = North American Spine Society Lumbar Spine Outcome Assessment Instrument; WOMAC = Western Ontario and McMaster Universities; SF-36 = Short Form-36; SES = Schmerzempfindungsskala (Pain Perception Questionnaire); HLC = Health Locus of Control; PCI = Proactive Coping Inventory; CSQ_R = revised Coping Strategies Questionnaire; SCQ = Self-administered Comorbidity Questionnaire

4. An exploratory analysis of predictor candidates of willingness to pay

Table 3. Predictors appearing in at least 7/10 imputed datasets in LASSO analysis with WTP as dependent

Health condition	WTP for	ICF component	Predictor	N	Exp(β_{med})*	Exp(β_{25})**	Exp(β_{75})**	N (imputed datasets including the predictor) ‡
Depression	expected effects	Body Functions	SF-36 pain (+)	104	1.0040	1.0035	1.0063	8/10
	perceived effects	Body Functions	SF-36 pain (+)	87	1.0049	1.0028	1.0083	8/10
	perceived effects	Body Functions	SF-36 vitality (+)	87	1.0162	1.0086	1.0232	7/10
	perceived effects	Activities & Participation	SF-36 physical function(+)	87	1.0045	1.0020	1.0052	9/10
Pain disorder	expected effects	Body Functions	SES affective pain sensation (-)	94	0.9740	0.9672	0.9859	8/10
	expected effects	Activities & Participation	SF-36 social function (+)	94	1.0051	1.0037	1.0061	7/10
Back pain	expected effects	Activities & Participation	SF-36 role-physical (+)	140	0.9954	0.9946	0.9961	10/10
	perceived effects	Activities & Participation	SF-36 role-physical (+)	109	0.9965	0.9954	0.9971	10/10
Osteoporosis	expected effects	Activities & Participation	SF-36 role-physical (+)	94	1.0008	1.0006	1.0022	7/10
Osteoarthritis	perceived effects	Activities & Participation	SF-36 role-emotional (+)	87	1.0024	1.0009	1.0032	7/10

* β_{med} =Median of parameter estimate over 10 imputations

** β_{25} , β_{75} =25th or 75th percentiles of parameter estimate over 10 imputations

‡ a total of 10 imputations were done to replace missing values

SF-36=Short Form-36; SES=Schmerzempfindungsskala (Pain Perception Scale)

(+) higher score indicates better health; (-) higher score indicates poorer health

4. An exploratory analysis of predictor candidates of willingness to pay

Table 4: Significant ($p < 0.05$) predictors of final GLM analysis with WTP as dependent

Health condition	WTP for	ICF component	Predictor	N	Exp (β)	95% CI	R^2_{final} (R^2_{base})*
Depression	perceived effects	Body Functions	ADS (-)	78	0.977	0.959; 0.996	43% (33%)
Pain disorder	expected effects	Personal Factors	School education	93	1.548	1.041; 2.302	42% (40%)
Back pain	perceived effects	Personal Factors	School education	93	1.621	1.030; 2.549	29% (15%)
	perceived effects	Personal Factors	CSQ pain decrease (+)	93	1.244	1.058; 1.462	29% (15%)
Osteoporosis	expected effects	Activities & Participation	HAQ (-)	83	0.723	0.539; 0.969	30% (24%)
	expected effects	Personal Factors	HLC internality (+)	83	1.034	1.007; 1.061	30% (24%)
	perceived effects	Activities & Participation	HAQ (-)	85	0.605	0.436; 0.841	26% (20%)
Osteoarthritis	expected effects	Personal Factors	School education	82	2.011	1.413; 2.863	33% (26%)
	perceived effects	Personal Factors	School education	73	1.894	1.240; 2.893	26% (17%)

* R^2_{base} , R^2_{final} correspond to Nagelkerke's R^2 of the base model (intercept, income, scope) and the final model (intercept, income, scope, predictors).

School education (ref=elementary); ADS=Allgemeine Depressionsskala (General Depression Scale); only used in depression); CSQ=Pain Coping Questionnaire; HAQ=Health Assessment Questionnaire (only used in osteoporosis); HLC=Health Locus of Control

(+) higher score represents either a higher effectiveness (CSQ pain decrease) or stronger beliefs (HLC internality); (-) higher score indicates poorer health

5 Conclusion

This doctoral thesis investigated factors beyond income and scope that predict the WTP for small to moderate health effects among musculoskeletal and psychosomatic patients undergoing rehabilitation interventions.

Eliciting the WTP based on contingent valuation surveys in which respondents answer questions about how they would act in hypothetical situations has never been a straightforward process. Many efforts have been made to develop valuation methodologies, however, the key issue remains as to whether hypothetical WTP would match the actual purchase decision [23]. In the absence of the gold standard (direct comparison of hypothetical statements and real behaviors), attempts are made to evaluate other aspects of validity, such as construct validity.

Assessing construct validity includes associations with factors derived from theory or empirical evidence predicting the WTP for health benefits. Theoretically, the WTP approach assumes that the valuation depends on both the characteristics of the good being valued and the characteristics of the individual. With regard to the “good” economic theory suggests that the WTP should be sensitive to the amount of the good or service supplied, this sensitivity is tested in scope tests. Concerning respondents’ characteristics, economic theory assumes that WTP should vary with ability to pay. Many studies therefore investigated the influence of income; a few studies addressed the scope effect. However, as empirical research has shown, income and scope cannot fully explain the WTP and there are other respondents’ characteristics affecting the WTP. This doctoral thesis contributes to the knowledge and understanding of predictors of WTP beyond income and scope. The objective

was therefore to investigate the influence of respondents' characteristics on WTP using the WTP for small to moderate health effects of rehabilitation interventions, as a case in point.

In the first part of this thesis predictors of WTP for changes in pain and related symptoms were identified on the basis of a systematic review. The review focused on pain because pain relief was the main outcome to be valued in the WTP study to which the subsequent predictor analysis referred. The predictors derived from the review included factors representing almost all of the ICF components. Predictors were related to functional impairments, limitations in activities and participation, personal and environmental factors. Moreover, aspects such as severity of illness, general health, satisfaction, and health expenditures were also covered. Only a few studies included a scope variable. The interpretation and understanding of predictors varied widely. The set of identified predictors can serve as the starting point for further research on the validity of WTP for health outcomes such as pain and related symptoms. In summarizing the available evidence from the existing literature of predictors of WTP for changes in pain and related symptoms, this review helps to spare difficult and time-consuming analyses of WTP-studies in order to find out what is known on respondents' characteristics affecting the WTP beyond income.

In the second part, an exploratory analysis of a comprehensive set of putative predictors, based on the findings of the systematic review, was conducted to find relevant factors predicting the WTP of patients with five different chronic health conditions undergoing rehabilitation interventions. Beyond income and scope, the analysis yielded important predictor candidates which ranged across the ICF

components. The predictor candidates were related to Body Functions, Activities and Participation, and Personal Factors. The predictor candidates were relevant and differed for different health conditions. Together, these results show that analyzing respondents' characteristics such as functional disability, activity limitations, coping and school education is useful for explaining WTP responses in the context of rehabilitation interventions. Since WTP responses based on contingent valuation are entirely hypothetical and criticized for being irrational and upwardly biased [140-142], comparing them to key aspects of the underlying health condition as well as socio-demographic features suggests that WTP responses are reasonable and credible.

The findings of this exploratory study provide tentative evidence for an influence of the identified predictors on WTP. Furthermore, this study permits preliminary recommendations on ways to identify predictors beyond income and scope when examining the WTP for the improvement of human functioning and minimization of disability using the rehabilitation strategy. Based on this study it is recommended to

- capture information about aspects of functioning relevant to the specific health condition (studies may rely to the ICF core sets which have been developed for specific chronic conditions),
- measure aspects of functioning with a generic health status instrument (e.g. SF-36 or WHODAS II) to increase comparability across studies (condition-specific measures of functioning may be used to address very specific issues of a health condition not covered by generic instruments)
- account for socio-demographic factors such as school education which are usually gathered in studies.

These recommendations may be useful to specify relevant predictors in future studies which should be encouraged to validate the WTP for health benefits, especially in the context of rehabilitation interventions. Analyzing aspects of functioning and personal factors, in addition to income and scope, enables a more comprehensive understanding of WTP responses and may help to reduce scepticism about applying this method in the economic evaluation of health care interventions.

6 Summary

Background

Economic issues involved in the evaluation of health and health care interventions are recognised in many countries to be of great importance. Although cost-effectiveness and cost-utility analysis have been the most widely used types of economic evaluation, there is growing interest in cost benefit analysis and the use of willingness to pay (WTP) as a measure of health benefits. One typical technique to elicit the WTP is contingent valuation (CV) which involves asking individuals directly in a hypothetical survey the maximum amount they are willing to pay to obtain a specific health improvement. Given that WTP figures derived from contingent valuations are based on responses to hypothetical questions, a crucial question is whether respondents can provide meaningful answers. Therefore contingent valuation surveys have to show that they obtain valid WTP responses. In the absence of the gold standard to directly compare hypothetical statements and real behaviors, attempts are made to evaluate other aspects of validity, such as construct validity. Assessing construct validity includes associations with predictors derived from theoretical or empirical evidence. Theoretically, the WTP approach assumes that the valuation depends on both the characteristics of the good being valued and the characteristics of the individual. According to economic theory, WTP values should increase with ability to pay and vary by the scope of the benefits arising from a health care program (so-called scope effect). Therefore, many studies investigated the influence of income on WTP; a few studies addressed the scope effect. However, as empirical research has been shown, income and scope cannot fully explain the WTP. Therefore it is important to determine additional relevant respondents' characteristics affecting the WTP for health improvements.

Objectives

The overall objective of this doctoral thesis was to explore respondents' characteristics beyond income and scope predicting the WTP for health benefits from contingent valuation studies. The specific aims were 1) to identify relevant predictors from other studies investigating the WTP for changes in pain and related symptoms, and 2) to examine putative predictors of WTP for small to moderate health effects of rehabilitation interventions.

With respect to these specific aims the doctoral thesis was subdivided into two parts. The first part presents predictor candidates of WTP based on a systematic review focusing on WTP for changes in pain and related symptoms. The review focused on this issue because pain relief was the main outcome to be valued in the WTP study to which the subsequent predictor analysis refers. The second part presents an exploratory analysis of predictors of WTP for small to moderate health effects among musculoskeletal and psychosomatic patients undergoing rehabilitation. A comprehensive set of predictor candidates including aspects of functioning and personal factors was used and the influence on WTP was compared across five different chronic health conditions (osteoarthritis, osteoporosis, back pain, pain disorder, and depression).

A systematic review of predictors of willingness to pay for changes in pain and related symptoms

The objective of this part of the doctoral thesis was to identify predictors of WTP beyond income, focusing on WTP for changes in pain and related symptoms.

For this purpose a systematic review of published WTP studies based on contingent valuation was carried out.

A literature search was conducted using electronic databases (MEDLINE, EMBASE, PsychInfo, Econlit) from 1966 to 2008. The International Classification of Functioning, Disability and Health (ICF) of the World Health Organization was used (as a tool) to classify the predictors of WTP.

21 studies were identified and reviewed. Beyond income there was a wide range of factors predicting the WTP for changes in pain and related symptoms. The identified predictors included factors covering almost all of the ICF components. Predictors were related to functional impairments, limitations in activities and participation, as well as personal and environmental factors. Moreover, aspects such as severity of illness, general health, satisfaction, and health expenditures were also covered. Only a few studies included a scope variable. The interpretation and understanding of predictors varied widely.

An exploratory analysis of predictor candidates of willingness to pay for small to moderate health effects of rehabilitation interventions

The objective of this part of the doctoral thesis was to explore respondents' characteristics beyond income and scope that predict the WTP based on contingent valuation for expected and perceived small to moderate health effects of rehabilitation interventions. The specific aims were (1) to examine a comprehensive set of putative predictors of WTP, and (2) to investigate whether they vary across health conditions.

Patients with five different health conditions (osteoarthritis, osteoporosis, back pain, pain disorder and depression) participated in an interview-based survey using CV questions to elicit the WTP for an expected and a perceived health effect. Additionally, patients completed several self-administered questionnaires measuring functioning and personal factors. All measures were classified using the International Classification of Functioning, Disability and Health (ICF) as a reference framework. Two statistical approaches were applied, each selecting predictors in different ways: the content-oriented stepwise-built generalized linear modelling (GLM), and the data-driven least absolute shrinkage and selection operator (LASSO).

The identified predictor candidates encompassed the ICF components Body Functions, Activities and Participation, as well as Personal Factors. This wide range of relevant predictors was not surprising, since it mirrored the complexity of rehabilitation medicine. The study showed that different predictors of WTP seemed to matter for different health conditions. While some predictors, including school education or limitations in daily activities as measured by the SF-36, were predictive in multiple health conditions, others affected the WTP in specific conditions. The two statistical methods returned two non-overlapping but, by themselves, reasonable sets of predictor candidates.

Conclusion

An exploratory analysis of a comprehensive set of putative predictors based on the findings of a systematic review was conducted to find relevant predictors of WTP for small to moderate health effects of rehabilitation interventions. Beyond

income and scope, the analysis yielded important predictor candidates which ranged across the ICF components. The predictor candidates were relevant and differed for different health conditions. Since WTP responses derived from contingent valuation are entirely hypothetical and criticized for being irrational and upwardly biased, they were compared with key aspects of the underlying health condition and also socio-demographic features suggesting that WTP responses are reasonable and credible. Although exploratory, the findings of this study provide tentative evidence of identified predictors of WTP and permit preliminary recommendations on ways to specify predictors of WTP beyond income and scope. Based on this study it is recommended to identify aspects of functioning relevant to a specific health condition, to use generic health status instruments to measure aspects of functioning and to consider school education, when analyzing predictors of WTP for the improvement of human functioning and minimization of disability using the rehabilitation strategy. The utilization of this strategy is recommended in future studies to validate the WTP for health benefits in the context of rehabilitation interventions.

7 Zusammenfassung

Hintergrund

Die Berücksichtigung ökonomischer Sachverhalte in der Evaluation der Gesundheitsversorgung genießt in vielen Ländern hohe Priorität. Obgleich Kosten-Effektivitäts-Analysen und Kosten-Nutzwert-Analysen die am weitesten verbreiteten Methoden der ökonomischen Evaluation sind, steigt das Interesse für Kosten-Nutzen Analysen und den Einsatz des Zahlungsbereitschaftsansatzes zur Messung des Gesundheitsnutzens. Die Zahlungsbereitschaft wird häufig mit Hilfe der kontingenten Bewertungsmethode erhoben, mit der Individuen in hypothetischen Szenarien direkt ihre maximale Zahlungsbereitschaft für eine bestimmte Gesundheitsverbesserung offenlegen. In Anbetracht des hypothetischen Charakters der angegebenen Zahlungsbereitschaften erhebt sich die entscheidende Frage, ob Individuen plausible und valide Antworten liefern. Dies führt zur Forderung, die Validität der mit der kontingenten Bewertungsmethode ermittelten Zahlungsbereitschaften zu belegen. Mangels eines Goldstandards, mit dessen Hilfe hypothetische Aussagen mit dem tatsächlichen Verhalten verglichen werden könnten, wird versucht andere Aspekte der Validität, wie die Konstruktvalidität, zu evaluieren. Die Bewertung der Konstruktvalidität erfolgt anhand der Beziehungen zwischen Zahlungsbereitschaft und Prädiktoren, die aus der ökonomischen Haushaltstheorie abgeleitet werden oder sich empirisch als bedeutend erwiesen haben. Theoretisch wird davon ausgegangen, dass die Zahlungsbereitschaft von den Eigenschaften des Bewertungsgegenstandes und den Merkmalen des Befragten abhängt. Entsprechend der ökonomischen Theorie sollte die Zahlungsbereitschaft mit der Höhe der Zahlungsfähigkeit und mit der Größe des Nutzengewinns aus einem Gesundheitsprogramm (sogenannter Scopeeffekt) steigen. Deshalb wurde der Einfluß des Einkommens auf die

Zahlungsbereitschaft in vielen Studien untersucht, nur wenige Studien beschäftigten sich mit dem Scopeeffekt. Wie die empirische Forschung jedoch gezeigt hat, kann die Zahlungsbereitschaft nicht vollständig über ökonomische Faktoren wie Einkommen und Scopeeffekt erklärt werden. Aus diesem Grund kommt der Bestimmung weiterer relevanter Merkmale der Befragten, die die Zahlungsbereitschaft für gesundheitliche Verbesserungen beeinflussen, erhebliche Bedeutung zu.

Ziele

Das allgemeine Ziel dieser Doktorarbeit besteht darin, neben dem Einkommen und der Scopevariable Merkmale der Befragten zu erforschen, die die Zahlungsbereitschaft für einen gesundheitlichen Nutzen in kontingenten Bewertungsstudien vorhersagen. Die spezifischen Ziele liegen darin a) relevante Prädiktoren aus anderen Studien zur Zahlungsbereitschaft für eine Veränderung der Schmerzsituation und assoziierter Symptome zu bestimmen und b) potentielle Prädiktoren im Kontext der Zahlungsbereitschaft für kleine und mittlere Gesundheitseffekte von rehabilitativen Maßnahmen zu untersuchen.

Unter Berücksichtigung dieser spezifischen Ziele gliedert sich die Doktorarbeit in zwei Teile: Im ersten Teil werden auf der Basis einer systematischen Literaturübersicht, die sich auf Studien der Zahlungsbereitschaft für eine Veränderung der Schmerzsituation und assoziierter Symptome bezieht, Prädiktoren bestimmt. Der Fokus der Literaturanalyse ergibt sich daraus, dass eine Schmerzreduktion der primäre Bewertungsgegenstand der Zahlungsbereitschaftsstudie ist, auf die sich die Prädiktorenanalyse bezieht. Der zweite Teil der Doktorarbeit besteht in einer explorativen Analyse von Prädiktoren der

Zahlungsbereitschaft für kleine und mittlere Gesundheitseffekte rehabilitativer Maßnahmen bei muskuloskelettalen und psychosomatischen Patienten. Eine Vielzahl von Prädiktorkandidaten, die Merkmale der Funktionsfähigkeit und Personenfaktoren umfassen, wird verwendet und der Einfluß auf die Zahlungsbereitschaft über fünf verschiedene chronische Erkrankungen (Osteoarthritis, Osteoporose, Rückenschmerz, somatoforme Schmerzstörungen, Depressionen) verglichen.

Eine systematische Literaturübersicht von Prädiktoren der Zahlungsbereitschaft für eine Veränderung der Schmerzsituation und assoziierter Symptome

Das Ziel dieses Teils der Doktorarbeit besteht darin, neben dem Einkommen Prädiktoren der Zahlungsbereitschaft zu bestimmen, wobei der Fokus auf der Zahlungsbereitschaft für eine Veränderung der Schmerzsituation und assoziierter Symptome lag. Zu diesem Zweck wurde eine systematische Literaturübersicht von veröffentlichten Zahlungsbereitschaftsstudien, die auf der kontingenten Bewertungsmethode basieren, durchgeführt.

Eine Literatursuche wurde in den elektronischen Datenbanken MEDLINE, EMBASE, PsychInfo und Econlit im Zeitraum von 1966 bis 2008 durchgeführt. Bei der Gruppierung der Prädiktoren kam die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) der Weltgesundheitsorganisation zur Anwendung.

21 Studien wurden identifiziert und ausgewertet. Neben dem Einkommen gab es eine Reihe verschiedener Prädiktoren der Zahlungsbereitschaft für eine Veränderung der Schmerzsituation und assoziierter Symptome. Die identifizierten Prädiktoren deckten die ICF Komponenten weitgehend ab. Die Prädiktoren verkörperten funktionale Schädigungen, Beeinträchtigungen von Aktivitäten und der Teilhabe, wie auch Personen- und Umweltfaktoren. Darüber hinaus deckten sie Schweregrad der Erkrankung, allgemeine Gesundheit, Zufriedenheit und Gesundheitsausgaben ab. Nur wenige Studien bezogen eine Scopevariable ein. Verständnis und Interpretation der Prädiktoren unterschieden sich stark.

Eine explorative Analyse von Prädiktorkandidaten der Zahlungsbereitschaft für kleine und mittlere Effekte rehabilitativer Maßnahmen

Das Ziel dieses Teils der Doktorarbeit ist es, neben dem Einkommen und der Scopevariable Merkmale der Befragten zu erforschen, die die Zahlungsbereitschaft basierend auf der kontingenten Bewertungsmethode für erwartete und erfahrene kleine und mittlere Gesundheitseffekte von rehabilitativen Maßnahmen vorhersagen. Die spezifischen Ziele bestehen darin, eine Vielzahl von potentiellen Prädiktoren simultan zu untersuchen und zu prüfen, ob sie von Krankheit zu Krankheit variieren.

Patienten mit fünf verschiedenen chronischen Erkrankungen (Osteoarthritis, Osteoporose, Rückenschmerzen, somatoforme Schmerzstörungen, Depressionen) nahmen an der Interview-basierten Erhebung der Zahlungsbereitschaft für erwartete und erfahrene Gesundheitseffekte teil. Die Zahlungsbereitschaft wurde mit der kontingenten Bewertungsmethode erhoben. Zusätzlich füllten die Patienten selbständig mehrere Fragebögen zur Messung der Funktionsfähigkeit und

persönlicher Faktoren aus. Alle erhobenen Merkmale wurden entsprechend der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) klassifiziert. Zwei statistische Verfahren, die auf unterschiedliche Weise Prädiktoren auswählen, wurden angewandt: die inhaltlich motivierte und schrittweise aufgebaute generalisierte lineare Modellierung und das rein Daten-basierte LASSO (least absolute shrinkage and selection operator), ein Verfahren zur simultanen Koeffizientenschätzung und Variablenselektion.

Die identifizierten Prädiktorkandidaten umfassten die ICF Komponenten Körperfunktionen, Aktivitäten und Teilhabe, wie auch Personenfaktoren. Die Bandbreite der Prädiktoren überraschte nicht in Anbetracht der Komplexität der Rehabilitationsmedizin. Die Studie zeigte, dass sich die Prädiktoren bezüglich der Erkrankungen anscheinend unterschieden. Während einige Prädiktoren wie Schulbildung oder Einschränkungen im täglichen Leben gemessen mit der SF-36 bei vielen Erkrankungen die Zahlungsbereitschaft vorhersagen konnten, beeinflussten andere wiederum die Zahlungsbereitschaft nur bei ganz spezifischen Erkrankungen. Die beiden statistischen Verfahren mündeten in zwei unterschiedliche jedoch in sich schlüssige Gruppen von Prädiktoren.

Schlussfolgerung

Zur Bestimmung relevanter Prädiktoren der Zahlungsbereitschaft für kleine und mittlere Effekte rehabilitativer Maßnahmen wurde eine explorative Analyse einer Vielzahl von Kandidaten, die aus einer systematischen Literaturübersicht abgeleitet wurden, durchgeführt. Neben dem Einkommen und der Scopevariable deckte die Analyse wichtige Prädiktorkandidaten auf, die sich über die verschiedenen ICF

Komponenten verteilt. Die Prädiktorkandidaten erwiesen sich als relevant und unterschieden sich hinsichtlich der Erkrankungen. Da Angaben zur Zahlungsbereitschaft, die auf dem kontingenten Bewertungsansatz basieren, als hypothetisch eingestuft und als irrational und verzerrt kritisiert werden, mag der Vergleich mit wichtigen Erkrankungsmerkmalen und sozio-demografischen Faktoren ihre Sinnhaftigkeit und Glaubwürdigkeit unterstützen. Trotz ihres explorativen Charakters geben die Ergebnisse dieser Studie vorläufige Hinweise auf die Evidenz der gefundenen Prädiktorkandidaten und erlauben erste Empfehlungen zum Vorgehen bei der Bestimmung von Prädiktoren der Zahlungsbereitschaft neben dem Einkommen und dem Scopeeffekt. Auf der Basis der vorliegenden Studie wird für die Analyse von Prädiktoren der Zahlungsbereitschaft für eine Verbesserung von Funktionsfähigkeit und Minimierung von Behinderung mittels rehabilitativer Maßnahmen empfohlen, die für eine Erkrankung relevanten Aspekte der Funktionsfähigkeit zu identifizieren, generische Instrumente zur Messung der Funktionsfähigkeit einzusetzen und die Schulbildung als Kriterium zu betrachten. Die Anwendung dieser Strategie wird in künftigen Studien zur Validierung der Zahlungsbereitschaft für einen gesundheitlichen Nutzen im Kontext rehabilitativer Maßnahmen empfohlen.

8 References

1. Blumenschein, K. and M. Johannesson, *Incorporating quality of life changes into economic evaluations of health care: an overview*. Health Policy, 1996. **36**(2): p. 155-166.
2. Olsen, J., R. Smith, and A. Harris, *Economic theory and the monetary valuation of health care: an overview of the issues as applied to the economic evaluation of health care programs*. Centre for Health Program Evaluation (CHPE) Working Paper No. 82, 1999.
3. Gafni, A., *Willingness to pay in the context of an economic evaluation of healthcare programs: theory and practice*. Am J Manag Care, 1997. **3** (Suppl): p. S21-32.
4. Drummond, M.F., et al., *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. 2005, New York: Oxford University Press.
5. Canadian Agency for Drugs and Technologies in Health, *Guidelines for the economic evaluation of health technologies: Canada*. 3rd ed. 2006, Ottawa: CADTH.
6. Ryan, M., V. Watson, and M. Amaya-Amaya, *Methodological issues in the monetary valuation of benefits in healthcare*. Expert Rev Pharmacoecon Outcomes Res, 2003. **3**(6): p. 717-27.
7. Johannesson, M. and B. Jonsson, *Economic evaluation in health care: is there a role for cost-benefit analysis?* Health Policy, 1991. **17**(1): p. 1-23.
8. Hammitt, J.K., *QALYs versus WTP*. Risk Analysis, 2002. **22**(5): p. 985-1001.
9. Rhoads, S.E., *The Economist's View of the World: Government, Markets, and Public Policy*. 1985, Cambridge: Cambridge University Press.
10. Özdemir, S., F.R. Johnson, and A.B. Hauber, *Hypothetical bias, cheap talk, and stated willingness to pay for health care*. Journal of Health Economics, 2009. **28**(4): p. 894-901.
11. Sach, T.H., R.D. Smith, and D.K. Whyne, *A 'league table' of contingent valuation results for pharmaceutical interventions: a hard pill to swallow?* Pharmacoeconomics, 2007. **25**(2): p. 107-127.
12. Sampietro Colom, L., V.L. Phillips, and A.B. Hutchinson, *Eliciting women's preferences in health care: a review of the literature*. Int J Technol Assess Health Care, 2004. **20**(2): p. 145-155.

13. Smith, R.D., *Construction of the contingent valuation market in health care: a critical assessment*. Health Economics, 2003. **12**(8): p. 609-28.
14. Ryan, M., et al., *Eliciting public preferences for healthcare: a systematic review of techniques*. Health Technology Assessment (Winchester, England), 2001. **5**(5): p. 1-186.
15. Olsen, J.A. and R.D. Smith, *Theory versus practice: a review of 'willingness-to-pay' in health and health care*. Health Economics, 2001. **10**(1): p. 39-52.
16. Klose, T., *The contingent valuation method in health care*. Health Policy, 1999. **47**(2): p. 97-123.
17. Bala, M.V., J.A. Mauskopf, and L.L. Wood, *Willingness to pay as a measure of health benefits*. Pharmacoeconomics, 1999. **15**(1): p. 9-18.
18. Diener, A., B. O'Brien, and A. Gafni, *Health care contingent valuation studies: a review and classification of the literature*. Health Econ, 1998. **7**(4): p. 313-26.
19. Bayoumi, A.M., *The measurement of contingent valuation for health economics*. Pharmacoeconomics, 2004. **22**(11): p. 691-700.
20. Liljas, B. and K. Blumenschein, *On hypothetical bias and calibration in cost-benefit studies*. Health Policy, 2000. **52**(1): p. 53-70.
21. Hammitt, J.K., *Valuing life-saving: is contingent valuation useful?* Risk in Perspective, 2003. **8**(3).
22. Olsen, J.A., C. Donaldson, and J. Pereira, *The insensitivity of willingness-to-pay to the size of the good: new evidence for health care*. Journal of Economic Psychology, 2004. **25**(4): p. 445-460.
23. Mitchell, R. and R. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method*. 1989, Washington, DC: Resources for the Future.
24. Kenkel, D., *On valuing morbidity, cost-effectiveness analysis, and being rude*. Journal of Health Economics, 1997. **16**(6): p. 749-757.
25. Gold, M.R., et al., *Cost-effectiveness in Health and Medicine*. 1996, New York, Oxford: Oxford University Press.
26. Whynes, D.K., E. Frew, and J.L. Wolstenholme, *A comparison of two methods for eliciting contingent valuations of colorectal cancer screening*. J Health Econ, 2003. **22**(4): p. 555-74.

27. Zillich, A.J., et al., *Assessment of the relationship between measures of disease severity, quality of life, and willingness to pay in asthma*. *Pharmacoeconomics*, 2002. **20**(4): p. 257-65.
28. Eastaugh, S.R., *Willingness to pay in treatment of bleeding disorders*. *Int J Technol Assess Health Care*, 2000. **16**(2): p. 706-10.
29. Donaldson, C., *Valuing the benefits of publicly-provided health care: does 'ability to pay' preclude the use of 'willingness to pay'?* *Soc Sci Med*, 1999. **49**(4): p. 551-63.
30. Appel, L.J., et al., *Risk reduction from low osmolality contrast media. What do patients think it is worth?* *Med Care*, 1990. **28**(4): p. 324-37.
31. Stavem, K., *Association of willingness to pay with severity of chronic obstructive pulmonary disease, health status and other preference measures*. *Int J Tuberc Lung Dis*, 2002. **6**(6): p. 542-9.
32. O'Brien, B. and J.L. Viramontes, *Willingness to pay: a valid and reliable measure of health state preference?* *Med Decis Making*, 1994. **14**(3): p. 289-297.
33. Ramsey, S.D., et al., *Willingness to pay for antihypertensive care: evidence from a staff-model HMO*. *Soc Sci Med*, 1997. **44**(12): p. 1911-7.
34. Masiye, F. and C. Rehnberg, *The economic value of an improved malaria treatment programme in Zambia: results from a contingent valuation survey*. *Malaria Journal*, 2005. **4**(60).
35. Stavem, K., *Willingness to pay: a feasible method for assessing treatment benefits in epilepsy?* *Seizure*, 1999. **8**(1): p. 14-19.
36. Johannesson, M., B. Jönsson, and G. Karlsson, *Outcome measurement in economic evaluation*. *Health Economics*, 1996. **5**: p. 279-96.
37. Arrow, K., et al., *Report of the NOAA Panel on Contingent Valuation*. *Federal Register*, 1993. **58**(10): p. 4601-4614.
38. Dranitsaris, G., *A pilot study to evaluate the feasibility of using willingness to pay as a measure of value in cancer supportive care: an assessment of amifostine cytoprotection*. *Support Care Cancer*, 1997. **5**(6): p. 489-99.
39. He, M., et al., *Willingness to pay for cataract surgery in rural Southern China*. *Ophthalmology*, 2007. **114**(3): p. 411-6.

40. Bosompra, K., et al., *Psychosocial factors associated with the public's willingness to pay for genetic testing for cancer risk: a structural equations model*. Health Educ Res, 2001. **16**(2): p. 157-72.
41. Weigl, M., et al., *Inpatient rehabilitation for hip or knee osteoarthritis: 2 year follow up study*. Ann Rheum Dis, 2004. **63**(4): p. 360-8.
42. Angst, F., et al., *Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities*. J Rheumatol, 2002. **29**(1): p. 131-8.
43. Angst, F., A. Aeschlimann, and G. Stucki, *Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities*. Arthritis Rheum, 2001. **45**(4): p. 384-91.
44. Bellamy, N., *Outcome measurement in osteoarthritis clinical trials*. J Rheumatol Suppl, 1995. **43**: p. 49-51.
45. Dieppe, P., *Therapeutic targets in osteoarthritis*. J Rheumatol Suppl, 1995. **43**: p. 136-9.
46. World Health Organization, *International Classification of Functioning, Disability and Health: ICF*. 2001, Geneva: WHO.
47. Stucki, G., A. Cieza, and J. Melvin, *The international classification of functioning, disability and health (ICF): A unifying model for the conceptual description of the rehabilitation strategy*. Journal of Rehabilitation Medicine, 2007. **39**(4): p. 279-285.
48. Bijlsma, J.W., *Patient centred outcomes in osteoarthritis*. Ann Rheum Dis, 2005. **64**(1): p. 1-2.
49. Felson, D.T., *Epidemiology of hip and knee osteoarthritis*. Epidemiol Rev, 1988. **10**(1): p. 1-28.
50. Dreinhöfer, K., et al., *ICF Core Sets for osteoarthritis*. Journal of Rehabilitation Medicine, 2004. **36**(Suppl 44): p. 75-80.
51. Guccione, A.A., et al., *The effects of specific medical conditions on the functional limitations of elders in the Framingham Study*. Am J Public Health, 1994. **84**(3): p. 351-8.
52. World Health Organization, *Guidelines for preclinical evaluation and clinical trials in osteoporosis*. 1998, Geneva: WHO.

53. WHO Technical Report Series No. 843, *Assessment of Fracture Risk and its Application to Screening for Postmenopausal Osteoporosis. Report of a WHO Study Group*. 1994, Geneva: World Health Organization.
54. Bonjour, J.P., P. Ammann, and R. Rizzoli, *Importance of preclinical studies in the development of drugs for treatment of osteoporosis: a review related to the 1998 WHO guidelines*. *Osteoporos Int.*, 1999. **9**(5): p. 379-93.
55. Cummings, S.R. and L.J. Melton, *Epidemiology and outcomes of osteoporotic fractures*. *Lancet*, 2002. **359**(9319): p. 1761-7.
56. Rahman, N. and K. Bhatia, *Impairments and disability associated with arthritis and osteoporosis*. Arthritis Series No. 4. 2007, Canberra: Australian Institute of Health and Welfare (AIHW).
57. Agency for Health Care Policy and Research, *Acute Low-back Problems in Adults. Clinical Practice Guideline Number 14*. 1994, Washington, DC: U.S. Government Printing Office.
58. Linton, S.J. and K. Hallden, *Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain*. *Clin J Pain*, 1998. **14**(3): p. 209-15.
59. Loeser, J.D. and E. Volinn, *Epidemiology of low back pain*. *Neurosurg Clin N Am*, 1991. **2**(4): p. 713-8.
60. Australian Acute Musculoskeletal Pain Guidelines Group, *Evidence-based Management of Acute Musculoskeletal Pain - a Guide for Clinicians*. 2004, Brisbane: Academic Press.
61. United States Bone and Joint Decade, *The Burden of Musculoskeletal Diseases in the United States. Prevalence, Societal and Economic Cost*. 2008, Rosemont, IL: American Academy of Orthopaedic Surgeons.
62. Koes, B.W., et al., *Clinical guidelines for the management of low back pain in primary care: an international comparison*. *Spine*, 2001. **26**(22): p. 2504-14.
63. Linton, S.J., *A Cognitive-Behavioral Approach to the Prevention of Chronic Back Pain*, in *Psychological Approaches to Pain Management: A Practitioner's Handbook*, D. Turk and R. Gatchel, Editors. 2002, New York: The Guilford Press.
64. Borkan, J., et al., *Advances in the field of low back pain in primary care: a report from the Fourth International Forum*. *Spine*, 2002. **27**(5): p. E128-E132.

65. Heneweer, H.M., et al., *Psychosocial variables in patients with (sub)acute low back pain: an inception cohort in primary care physical therapy in the Netherlands*. Spine, 2007. **32**(5): p. 586-92.
66. Gatchel, R.J. and J. Dersh, *Psychological Disorders and Chronic Pain: Are There Cause- and -Effect Relationships?* in *Psychological Approaches to Pain Management: A Practitioner's Handbook*, D. Turk and R. Gatchel, Editors. 2002, New York: The Guilford Press.
67. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders. DSM-IV (4th ed)*. 1994, Washington, DC: American Psychiatric Association.
68. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders. DSM-III (3rd ed)*. 1980, Washington, DC: American Psychiatric Association.
69. Fröhlich, C., F. Jacobi, and H.-U. Wittchen, *DSM-IV pain disorder in the general population*. European Archives of Psychiatry & Clinical Neuroscience, 2006. **256**(3): p. 187-196.
70. World Health Organization, *World Health Report 2001. Mental Health: New Understanding, New Hope*. 2001, Geneva: WHO.
71. Cieza, A., et al., *ICF Core Sets for depression*. Journal of Rehabilitation Medicine, 2004. **36**: p. 128-134.
72. Chestnut, L.G., et al., *Measuring heart patients' willingness to pay for changes in angina symptoms*. Med Decis Making, 1996. **16**(1): p. 65-77.
73. Stucki, G., T. Ewert, and A. Cieza, *Value and application of the ICF in rehabilitation medicine*. Disability and Rehabilitation, 2002. **24**(17): p. 932-938.
74. Geyh, S. and A.e.a. Cieza, *Content comparison of health-related quality of life measures used in stroke based on the international classification of functioning, disability and health (ICF): a systematic review*. Qual Life Res, 2007.
75. Cieza, A., et al., *ICF linking rules: an update based on lessons learned*. J Rehabil Med, 2005. **37**(4): p. 212-18.
76. Allen, T.K., et al., *How much are patients willing to pay to avoid postoperative muscle pain associated with succinylcholine?* J Clin Anesth, 2007. **19**(8): p. 601-608.
77. Atchison, K.A., et al., *Baseline characteristics and treatment preferences of oral surgery patients*. J Oral Maxillofac Surg, 2007. **65**(12): p. 2430-2437.

-
78. Bala, M.V., et al., *Valuing outcomes in health care: a comparison of willingness to pay and quality-adjusted life-years*. J Clin Epidemiol, 1998. **51**(8): p. 667-76.
 79. Boonen, A., et al., *Willingness of patients with ankylosing spondylitis to pay for inpatient treatment is influenced by the treatment environment and expectations of improvement*. Ann Rheum Dis, 2005. **64**(11): p. 1650-1652.
 80. Byrne, M.M., K.J. O'Malley, and E. Suarez Almazor, *Ethnic differences in health preferences: analysis using willingness-to-pay*. J Rheumatol, 2004. **31**(9): p. 1811-1818.
 81. Cross, M.J., et al., *Determinants of willingness to pay for hip and knee joint replacement surgery for osteoarthritis*. Rheumatology, 2000. **39**(11): p. 1242-48.
 82. Donaldson, C. and P. Shackley, *Does "process utility" exist? A case study of willingness to pay for laparoscopic cholecystectomy*. Soc Sci Med, 1997. **44**(5): p. 699-707.
 83. Dranitsaris, G., *Pamidronate for the prevention of skeletal-related events in multiple myeloma. What does the public think it is worth?* Int J Technol Assess Health Care, 1999. **15**(1): p. 108-22.
 84. Ethgen, O., et al., *Do utility values and willingness to pay suitably reflect health outcome in hip and knee osteoarthritis? A comparative analysis with the WOMAC Index*. J Rheumatol, 2003. **30**(11): p. 2452-9.
 85. Fautrel, B., et al., *Valuing a hypothetical cure for rheumatoid arthritis using the contingent valuation methodology: the patient perspective*. J Rheumatol, 2005. **32**(3): p. 443-53.
 86. Greenberg, D., et al., *Willingness to pay for avoiding coronary restenosis and repeat revascularization: results from a contingent valuation study*. Health Policy, 2004. **70**(2): p. 207-16.
 87. Hamelsky, S.W., R.B. Lipton, and W.F. Stewart, *An assessment of the burden of migraine using the willingness to pay model*. Cephalalgia, 2005. **25**(2): p. 87-100.
 88. Jacobs, R.J., R.J. Moleski, and A.S. Meyerhoff, *Valuation of symptomatic hepatitis a in adults: estimates based on time trade-off and willingness-to-pay measurement*. Pharmacoeconomics, 2002. **20**(11): p. 739-47.
 89. Kartman, B., F. Andersson, and M. Johannesson, *Willingness to pay for reductions in angina pectoris attacks*. Med Decis Making, 1996. **16**(3): p. 248-53.

90. King Jr, J.T., et al., *Preference-based quality of life measurement in patients with cervical spondylotic myelopathy*. Spine, 2004. **29**(11): p. 1271-1280.
91. Narbro, K. and L. Sjostrom, *Willingness to pay for obesity treatment*. Int J Technol Assess Health Care, 2000. **16**(1): p. 50-9.
92. Ritchie, L.D. and N. Johnson, *A willingness-to-pay analysis: comparison of the patient preference for topical flurbiprofen local action transcutaneous patches and piroxicam gel*. Br J Med Econ, 1996. **10**(3): p. 291-302.
93. Thompson, M.S., *Willingness to pay and accept risks to cure chronic disease*. Am J Public Health, 1986. **76**(4): p. 392-6.
94. van-den-Bosch, J.E., et al., *Effect of postoperative experiences on willingness to pay to avoid postoperative pain, nausea, and vomiting*. Anesthesiology, 2006. **104**(5): p. 1033-9.
95. Werner, P. and I. Vered, *Women's willingness to pay out-of-pocket for drug treatment for osteoporosis before and after the enactment of regulations providing public funding: evidence from a natural experiment in Israel*. Osteoporos Int, 2002. **13**(3): p. 228-34.
96. Stucki, G., et al., *ICF Core Sets for rheumatoid arthritis*. J Rehabil Med, 2004(Suppl 44): p. 87-93.
97. Lundberg, L., et al., *Quality of life, health-state utilities and willingness to pay in patients with psoriasis and atopic eczema*. Br J Dermatol, 1999. **141**(6): p. 1067-75.
98. Zethraeus, N., et al., *The impact of hormone replacement therapy on quality of life and willingness to pay*. Br J Obstet Gynaecol, 1997. **104**(10): p. 1191-5.
99. van Houtven, G., et al., *Valuing avoided morbidity using meta-regression analysis: what can health status measures and QALYs tell us about WTP?* Health Economics, 2006. **15**(8): p. 775-795.
100. van Houtven, G., et al., *Valuation of morbidity losses: meta-analysis of willingness-to-pay and health status measures*. 2003, Final Report, Prepared for the Food and Drug Administration by Research Triangle Institute.
101. Schiffner, R., et al., *Willingness to pay and time trade-off: useful utility indicators for the assessment of quality of life and patient satisfaction in patients with port wine stains*. British Journal of Dermatology, 2002. **146**(3): p. 440-447.

102. Weaver, M., D.L. Patrick, and L.E.e.a. Markson, *Issues in the measurement of satisfaction with treatment*. The American Journal of Managed Care, 1997. **3**(4): p. 579-594.
103. Rheingans, R.D., et al., *Willingness to pay for prevention and treatment of lymphatic filariasis in Leogane, Haiti*. Filaria J, 2004. **3**(2).
104. Wagner, T.H., et al., *Does willingness to pay vary by race/ethnicity? An analysis using mammography among low-income women*. Health Policy, 2001. **58**(3): p. 275-88.
105. Fautrel, B., et al., *Willingness-to-pay for the cure of a chronic disease: Differences of perception between patients and the general population relative to rheumatoid arthritis*. J-Econ-Med, 2005. **23**(6): p. 329-350.
106. Bullinger, M., *German translation and psychometric testing of the SF-36 Health Survey: preliminary results from the IQOLA project*. Social Science & Medicine, 1995. **41**(10): p. 1359-66.
107. Ware, J.E., Jr. and C.D. Sherbourne, *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection*. Med Care, 1992. **30**(6): p. 473-83.
108. Sangha, O., et al., *The Self-administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research*. Arthritis Care & Research, 2003. **49**(2): p. 156-163.
109. Stucki, G., et al., *Evaluation of a German version of WOMAC (Western Ontario and McMasterUniversities) Arthrosis Index*. Z Rheumatol, 1996. **55**(1): p. 40-49.
110. Bellamy, N., et al., *Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee*. J Rheumatol, 1988. **15**(12): p. 1833-40.
111. Brühlmann, P., G. Stucki, and B. Michel, *Evaluation of a German version of the physical dimensions of the Health Assessment Questionnaire in patients with rheumatoid arthritis*. J Rheumatol, 1994. **21**(7): p. 1245-9.
112. Fries, J., et al., *Measurement of patient outcome in arthritis*. Arthritis & Rheumatism, 1980. **23**(2): p. 137-145.
113. Pose, B., et al., *Validation of the North American Spine Society Instrument for assessment of health status in patients with chronic backache*. Z Orthop Ihre Grenzgeb, 1999. **137**(5): p. 437-441.

114. Daltroy, L., et al., *The North American Spine Society Lumbar Spine Outcome Assessment instrument: reliability and validity tests*. Spine, 1996. **21**(6): p. 741-748.
115. Hautzinger, M. and M. Bailer, eds. *Allgemeine Depressions Skala*. 1993, Göttingen: Beltz Test GmbH.
116. Radloff, L.S., *The CES-D Scale: a self-report depression scale for research in the general population*. Applied Psychological Measurement, 1977. **1**(3): p. 385-401.
117. Geissner, E., *The Pain Perception Scale - a differentiated and change-sensitive scale for assessing chronic and acute pain*. Rehabilitation, 1995. **34**(4): p. XXXV-XLIV.
118. Melzack, R., *The McGill Pain Questionnaire: major properties and scoring methods*. Pain, 1975. **1**(3): p. 277-299.
119. Lohaus, A. and G. Schmitt, eds. *Fragebogen zur Erhebung von Kontrollüberzeugungen zu Krankheit und Gesundheit (KKG)*. 1989, Göttingen: Hogrefe.
120. Wallston, K., B. Wallston, and R. DeVellis, *Development of the Multidimensional Health Locus of Control (MHLC) Scales*. Health Education Monographs, 1978. **6**: p. 160-170.
121. Greenglass, E., R. Schwarzer, and S. Taubert. *The Proactive Coping Inventory (PCI): a multidimensional research instrument*. 1999 [cited 2009 June]; available from: <http://userpage.fu-berlin.de/~health/poland.htm>.
122. Riley, J. and M. Robinson, *CSQ: Five factors or fiction?* Clinical Journal of Pain, 1997. **13**(2): p. 156-162.
123. Robinson, M., et al., *The Coping Strategies Questionnaire: a large sample, item level factor analysis*. Clinical Journal of Pain, 1997. **13**(1): p. 43-49.
124. Tibshirani, R., *Regression shrinkage and selection via lasso*. Royal. Statist. Soc B., 1996. **58**(1): p. 267-288.
125. Honaker, J., G. King, and M. Blackwell. *Amelia II: A Program for Missing Data*. 2009 [cited 2009 April]; available from: <http://gking.harvard.edu/amelia/>.
126. O'Brien, R.M., *A caution regarding rules of thumb for variance inflation factors*. Quality & Quantity, 2007. **41**(5): p. 673-690.

127. Bickenbach, J.E., et al., *Models of disablement, universalism and the international classification of impairments, disabilities and handicaps*. Soc Sci Med, 1999. **48**(9): p. 1173-87.
128. Strine, T.W., et al., *Health-related quality of life, health risk behaviors, and disability among adults with pain-related activity difficulty*. American Journal of Public Health, 2005. **95**(11): p. 2042-48.
129. Mercado, A., L. Carroll, and D. Cassidy, *Coping with neck and low back pain in the general population*. Health Psychology, 2000. **19**(4): p. 333-338.
130. Main, C.J. and A.C. Williams, *Musculoskeletal pain*. BMJ, 2002. **325**(7363): p. 534-7.
131. Angst, F., et al., *Interdisciplinary rehabilitation in fibromyalgia and chronic back pain: a prospective outcome study*. J Pain, 2006. **7**(11): p. 807-15.
132. Stucki, G. and G. Grimby, *ICF core sets for chronic conditions*. J Rehabil Med, 2004. **36**(Suppl 44): p. 1-141.
133. World Health Organization. *World Health Organization Disability Assessment Schedule II (WHODAS II)*. 2001 [cited 2009 June]; available from: <http://www.who.int/icidh/whodas/>.
134. Dworkin, R.H., et al., *Core outcome measures for chronic pain clinical trials: IMMPACT recommendations*. Pain, 2005. **113**(1-2): p. 9-19.
135. Hackl, F. and G.J. Pruckner, *On the gap between payment card and closed-ended CVM-answers*. Applied Economics, 1999. **31**(6): p. 733-42.
136. Cameron, T.A. and D.D. Huppert, *OLS versus ML estimation of non-market resource values with payment card interval data*. Journal of Environmental Economics and Management, 1989. **17**(3): p. 230-46.
137. Turner, R.R., et al., *Patient-reported outcomes: instrument development and selection issues*. Value in Health, 2007. **10**: p. S86-93.
138. Stanton, J.M., et al., *Issues and strategies for reducing the length of self-report scales*. Personnel Psychology, 2002. **55**(1): p. 167-94.
139. Sloan, J.A., et al., *Analysis and interpretation of results based on patient-reported outcomes*. Value Health, 2007. **10** (Suppl 2): p. S106-15.
140. Kotchen, M.J. and S.D. Reiling, *Environmental attitudes, motivations, and contingent valuation of nonuse values: a case study involving endangered species*. Ecological Economics, 2000. **32**(1): p. 93-107.

141. Diamond, P.A. and J.A. Hausman, *Contingent valuation: is some number better than no number?* Journal of Economic Perspectives, 1994. **8**(4): p. 45-64.
142. Hausman, J.A., *Contingent Valuation: A Critical Assessment*. 1993, Amsterdam: Elsevier.

9 Appendix

Appendix 1: Search strategy

The full search strategy limited to titles and abstracts was as follows: (('willingness to pay' OR 'contingent valuation') AND ('health*' OR 'ill*' OR 'disease*' OR 'treatment' OR 'therapy' OR 'quality of life' OR 'testing' OR 'patients' OR 'nurs*' OR 'rehab*' OR 'functioning' OR 'participation' OR 'prevention' OR 'pain' OR 'Analgesic*') AND ('trial' OR 'study' OR 'survey') AND ('NOT cost effectiveness' OR 'NOT acceptability curve') AND ('predict*' OR 'determinant*' OR 'characteristic*' OR 'indicator*' OR 'associat*' OR 'correlat*' OR 'relat*' OR 'influenc*' OR 'affect*')).

Appendix 2: Explanations for significant predictors

Explanations of significant ($p \leq 0.05$) predictors of willingness to pay (WTP) which were **linked** to the ICF

ICF COMPONENT <i>ICF Category</i> 2nd level ICF category	Description of predictor	Interpretations provided by the authors
BODY FUNCTIONS		
Body functions	mJOA	A higher WTP being the equivalent of lower QOL (King)
b1 <i>Mental functions</i>		
b152 Emotional functions	CES_D	No explanation (Thompson)
	Anxiety state	No explanation (van den Bosch)
b2 <i>Sensory functions and pain</i>		
b280 Sensation of pain	Womac pain	Pain relief being the dominant determinant of WTP for joint replacement (Cross)
	Pain of unknown duration	No explanation (Hamelsky)
b4 <i>Functions of the cardiovascular, haematological, immunological and respiratory systems</i>		
b410 Heart functions	Angina pectoris attacks	Showing the validity of WTP estimates (Kartman)
ACTIVITIES & PARTICIPATION		
Activities & Participation	MIDAS	Moderate (MIDAS Grade III) not severe disability (MIDAS Grade IV) influenced the WTP. The reason for this may be that patients with severe disability are more likely to be

ICF COMPONENT <i>ICF Category</i> 2nd level ICF category	Description of predictor	Interpretations provided by the authors
d4 Mobility d450 Walking	HAQ	depressed and may have given up on the possibility of excellent relief (Hamelsky) It seems that people contemplating paying money for arthritis cure ask themselves how they would improve functionally (Thompson)
	50-foot walk	No explanation (Thompson)
ENVIRONMENTAL FACTORS e5 Services, systems and policies e580 Health services, systems and policies	Care in public clinic	Probably in large part a measure of ability to pay (Fautrel)
PERSONAL FACTORS¹⁾ Coping Experiences	Daily coping	No explanation (Hamelsky)
	Prior shingles	No explanation (Bala)
	Previous bypass	Patients having undergone successful bypass surgery in the past are less concerned about the possibility of a second operation compared with patients who have never experienced this procedure (Greenberg)
	Previous bypass	No explanation (Chestnut)
Knowledge	Knowledge about osteoporosis	Responsible for more rational decision making (Werner)
Sociodemographics Age		No explanation (Byrne, Kartman, Thompson) The higher WTP of older patients undergoing TKR may be related to the overall good health of older respondents (Cross)

ICF COMPONENT	Description of predictor	Interpretations provided by the authors
<i>ICF Category</i> 2nd level ICF category		
Education		The lower WTP for elderly people might be related to a lower likelihood of experiencing benefit from a private program because their children are old enough to be economically independent. It is also possible that age effect represents a lower ability to pay (Fautrel)
		Lower age at onset of obesity may represent the degree of sickness (Narbro)
		It is unclear why people who attended graduate school (a surrogate for high income) would be less likely to pay high amounts of money for relief (Hamelsky)
Gender		No explanation (Narbro)
Household size		No explanation (Bala)
		The positive effects of children on WTP may be related to the general desire of Canadian society to protect the dependents of patients with terminal diseases (Dranitsaris)
		No explanation (Narbro)
Health insurance		No explanation (Cross)
		Private health or drug insurance is probably a measure of ability to pay. However, to the extent that it represents a choice (of a job with insurance or to purchase health insurance), the effects of the insurance variables represent preferences of health versus other goods (Fautrel)
Race		Ethnic groups may place different values on difficulty in walking and mobility... Alternatively, it is possible that the reduced value for health improvement in African Americans, and to a lesser degree Hispanics, reflects a

ICF COMPONENT	Description of predictor	Interpretations provided by the authors
ICF Category 2nd level ICF category		
		lack of awareness of, or experience with, the benefits of treatment, perhaps due to past discrimination and reduced access to medical care (Byrne)
Religion		No explanation (Dranitsaris)
Social class		No explanation (Donaldson)

¹) Currently not classified in the ICF

mJOA=Western modification of the Japanese Orthopaedic Association scale; CES_D=Center for Epidemiological Studies Depression Scale;

WOMAC=Western Ontario and McMaster Universities; MIDAS=Migraine Disability Assessment Score; HAQ=Health Assessment Questionnaire

Explanations of significant ($p \leq 0.05$) predictors of willingness to pay (WTP) which were **not linked** to the ICF

Domains	Description of predictor	Interpretations provided by the authors
Health condition	Angina pectoris status	Indicating that WTP increased with the size of health improvement (Kartman)
	body weight	Current weight of obese people may represent the severity of illness (Narbro)
General health Expenditures (health related)	RAND current health	no explanation (Thompson)
	Prescription payment	It is logical that subjects who pay for prescription migraine treatment were more likely to be willing to pay than those who do not use prescription treatment (Hamelsky)
Satisfaction with healthcare services	Recommend replacement	Suggesting the benefits of joint replacement (Cross)
Scope	Risk reduction	Assuming risk reduction is a desired good, theoretical considerations would suggest that WTP should increase in proportion to the magnitude of risk reduction being offered (Greenberg)
	Reduction in angina attack rate	Showing the validity of WTP estimates (Kartman)
Miscellaneous	Bid	No explanation (Bala, Ethgen, Kartman)
		Reflecting starting point bias in iterative bidding procedures (Chestnut)
		Confirming internal validity of the WTP method elicited as dichotomous choice question (Greenberg)
	Group	No explanation (van Bosch)
	WTP before regulations	May represent a higher perceived risk of restenosis (Greenberg) The relationship between the WTP before and after the enactment of regulations was suggested to indicate the stability of WTP measures over time and therefore to support their validity (Werner)

Appendix 3: Content of scales and meaning of scores

The Short Form-36 (SF-36) German version contains eight multi-item scales: physical functioning, role limitation due to physical problems (role-physical), bodily pain, vitality, social functioning, role limitation due to emotional problems (role-emotional), mental health, general health perception. Higher scores represent better functioning. Additionally, there is a single item that provides an indication of perceived change in health as measured on a five-point Likert response scale of excellent, very good, good, fair, poor [106, 107].

The Self-administered Comorbidity Questionnaire (SCQ) contains items relating to 12 defined medical problems. A higher score indicates a greater burden of coexisting health conditions [108].

The German version of the Western Ontario and McMaster Universities (WOMAC) Osteoarthritis Index includes three subscales that cover the dimensions of pain (five items), stiffness (two items), and physical function (17 items). Lower scores reflect better function, less pain, or less stiffness [109, 110].

The German version of the disability dimension of the Health Assessment Questionnaire (HAQ_DI) is comprised of 20 questions in eight categories of daily activities (dressing, rising, eating, walking, hygiene, reach, grip, and usual activities). It yields a single index that indicates the extent of respondents' functional limitations. A higher score indicates more limitations [111, 112].

The German version of the lumbar North American Spine Society (NASS) questionnaire consists of 17 questions, 11 concerning pain and disability and, 6 about neurogenic symptoms. Lower scores on the two subscales represent better functioning [113, 114].

The German version of the Centre for Epidemiological Studies Depression Scale (CES_D), the ADS (Allgemeine Depressionsskala) contains 20 items measuring depressive symptoms [115, 116] with the higher scores indicating more symptoms.

The Pain Perception Scale (Schmerzempfindungsskala – SES) consists of 24 descriptors: 14 affective and 10 sensory pain sensations. A higher score indicates ascending intensity of pain [117]. The Pain Perception Scale is a continuation of the McGill Pain Questionnaire [118].

“Fragebogen zur Erhebung von Kontrollüberzeugungen zu Krankheit und Gesundheit“ (KKG) by Lohaus and Schmitt [119], the German version of the Health Locus of Control (HLC) scales by Wallston KA, Wallston BS, DeVellis R [120] was used for this thesis. The KKG consists of three health and illness-related control beliefs: internal health locus of control, powerful others health locus of control, and chance health locus of control. Higher scores indicate stronger beliefs that one`s outcome is directly the result of one`s behavior (internal), or that one`s outcome is under the control of powerful others or is determined by fate, luck or chance.

Two subscales of the German test version of the Proactive Coping Inventory (PCI), strategic planning and preventive coping, contain four and ten items

respectively. Individuals scoring high on the two subscales are seen as being able to take preparatory steps in coping with anticipated stress [121].

The CSQ_R is the short version of the Coping Strategies Questionnaire (CSQ) retaining 27 (items) of the original scales. They assess six cognitive pain coping techniques (distraction, catastrophizing, ignoring pain, distancing from pain, coping self-statements and praying). Additionally, there are two separate effectiveness (single) items, the perceived control over pain and the ability to decrease pain based on the use of the endorsed strategies. A higher score of the subscales reflects more frequent use of specific coping strategies and a higher score of the single items represents a higher effectiveness [122, 123].