Outcome measurement in hand surgery – a brief overview

M. Vlach1,2
1 Hand and Plastic Surgery Institute, Vysoké nad Jizerou, Czech Republic
2 Department of Orthopaedics, Second Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic

Summary

Many objective examinations can be used to evaluate the results of surgical care. In addition, a variety of specific questionnaires, so-called PROMs (Patient-reported Outcome Measures), can be used to assess a patient’s subjective perception. In recent years, PROMs, which assess the main areas of patients’ interest, have been considered an important element in the comprehensive measurement of treatment outcomes. Recently, this approach has also proven valuable for the possibility of evaluating the results of health care without the need for the physical presence of the patient in the medical facility. In the presented overview, a brief list of objective measurements in the area of the hand and wrist is provided. It then focuses on the most used PROMs, together with a summary of the basic theory of their use.

Key words

outcome – measurement – evaluation – questionnaire – PROMs – hand surgery

Introduction

Objective measurement of results is an essential part of a comprehensive evaluation of surgical procedures [1]. The author conducted a literature search and compiled an overview of methods for evaluating the results usable for the needs of hand surgery. The first part contains a list of relevant objective measurements in the area of the hand and wrist. A list of several most used standardised questionnaires (PROMs) follows, together with the basic theory of the use of the given tools.

Why should we measure?

The need to define and evaluate results in (not only) hand surgery arises from several sources, mainly from the need of surgeons and therapists to evaluate their own results and to be able to compare them with others in the literature.

Tools and tests used must be, among others, independent and resistant to modifications by patients and those who evaluate the results for (employment and) legal reasons [2]. Health care outcomes cannot and must not be considered in isolation without considering the many factors that affect the end result. These are input factors such as the patient and their personality and health status, the condition of the limb before surgery, as well as the nature and severity of the injury or illness. In addition, procedural aspects must be considered, such as the characteristics of the surgical procedure itself and its complexity, as well as the presence or absence of complications. Furthermore, the output factors, which include the level of expertise and experience of the facility providing care, are assessed.

At the same time, accurate evaluation of the results of acute and elective care is the first condition for comparing the results of individual procedures or workplaces among each other and for using the achieved results to create publications and provide conditions for the progress of knowledge of the studied issues.

In the Czech Republic, the use of standardised questionnaires for outcome evaluation in hand surgery is not widely spread yet, presumably at least partially due to the existence of only a minimum of official Czech translations. The creation and validation of Czech versions of individual PROMs, together with the proper methodology of their use, is room space for future studies.

To compile this review, the Medline database was searched for articles from the period 2010–2021 based on the keywords “PROMs AND Hand Surgery”. A total of 128 articles were found from which articles that did not deal with the issue were discarded. Additional literary sources were found by searching the bibliography of essential articles on the subject. A total of 24 articles dealing with the given topic were selected and PROMs mentioned in them were reviewed.

Evaluation of the topic

Types of tests

Tests used for outcome measurement in hand surgery, whether site-specific or...
condition-specific, can be divided into instrumental and clinical tests and questionnaires. Questionnaires can be further divided into smaller groups: questionnaires specific to the location, to a condition, or generic questionnaires (e.g., quality of life assessment). From some studies comparing the yield of different questionnaires among each other, it is possible to deduce the suitability of using more than one questionnaire at once [3,4].

**What to measure?**

In modern medicine, there is an effort to perform examinations using objective procedures. However, this effort may encounter the fact that some qualities are difficult to objectify, especially those related to sensory functions, or function (of the hand) in general.

The ranges of joint movement, measured in degrees, are probably easiest to objectively evaluate. Here, too, the question arises of the magnitude of the force applied in order to measure passive range of motion. In the 1980s, Brand promoted so-called TACs, or torque-angle-curves, for measuring the angles of range of motion in joints, based on a precisely determined force (Fig. 1) [5]. However, this technique has not found greater application in practice. The extent of the flexion of the fingers, or more precisely their stiffness, can be additionally evaluated using the so-called pulp-to-palm distance. Even here, however, it is necessary to pay attention to the accuracy of measurement in accordance with valid definitions of “pulp” and “palm” (Fig. 2).

The force of compression (both grip and pinch strength) is another easily objectively measurable quality. For both qualities, gauges are available, as well as established procedures to detect submaximal effort of the test subject, e.g., repeated testing.

The evaluation of sensual perception is more difficult to objectify. There are relatively objective methods, such as the evaluation of two-point discrimination (Fig. 3) or the use of Semmes-Weinstein monofilaments. Other tests that have not been used in wider practice include, for example, the measurement of skin conductivity and resistance, which can be used in the evaluation of nerve injuries and regeneration, and it even allows the formulation of a prognosis of the health condition.

Acute pain is the last often objectively stated quality. For its evaluation, a number of scales have been developed over the years, with an 11-point Visual Analogue Scale (VAS score) used in adults and Faces Pain Scale (-Revised) used in children being probably the most often used ones [6]. However, even in this case, the correlation between groups of patients is problematic due to psychological, cultural and emotional differences [7]. Chronic pain constitutes an even larger problem with regards to the evaluation of results [8].

In 1963, Buchwald introduced the term “activities of daily living” (ADL). These activities include mobility, the ability to get dressed and undressed, use the toilet and perform basic personal hygiene, eat and drink, communicate, maintain housekeeping, but also employ in areas such as work and recreation [9]. Although the desire to evaluate the ability to perform these activities in an individual may seem unscientific, it plays an important role in the evaluation of the outcome as such, from the point...
of view of both the patient and the surgeon or the therapist [2].

A number of tests have been developed to assess the work ability of patients with hand disabilities. Nevertheless, a mere effort to list them would go beyond the scope of this paper. In the Czech Republic, they are performed predominantly by occupational therapists. They are subject to the same requirements that are valid for the individual tests described below.

Methods of evaluation of results, which cover the main areas of patient's interest, are considered an important element of comprehensive measurement of the results of surgical procedures [3]. In any evaluation, one should not forget Young’s words: “It is a common misconception that all tests that generate numbers are necessarily objective” [10]. Conversely, the evaluation of areas such as ADL should not be marginalised simply because it does not generate numbers, and care must be taken to apply a quantitative approach [2].

**How to measure?**

Regarding tools for measuring results, it is desirable to assess several key criteria. These are validity, reliability and sensitivity. Furthermore, it is appropriate to put accent on the interpretability of the results obtained, and, as recent experience has shown, resistance to unexpected circumstances, such as a pandemic, has also become an increasingly important factor.

Validity describes how accurately an instrument measures what is being measured. It is usually divided into two components: content validity and criterion validity. The first assesses whether the test measures what is to be measured, the latter indicates the degree of compliance of the results with a set criterion. Criterion validity is further divided into concurrent validity (i.e. whether the results obtained correlate with the results of another recognized evaluation method), and predictive validity (which determines the ability of the test to predict the result in the future).

Reliability is the ability to measure a given quantity in a reproducible way. It also has two components. The first is reproducibility; the second is the so-called internal consistency, i.e. the assumption that all items measuring one property yield sufficiently high correlations.

Sensitivity refers to the ability to evaluate a relatively small change in a measured variable.

Finally, interpretability tells us whether both the examiner and the test subject understand the obtained results.

The demand for the methods to be usable even under the conditions of a pandemic has arisen mainly in recent years in reaction to the COVID-19 pandemic, which has significantly altered the functioning of health care worldwide (and, as a side effect, the possibilities to evaluate results) [11]. The questionnaire survey can be performed remotely, either by physical correspondence or by electronic communication methods, such as an e-mail or a web application. When trying to obtain data by correspondence, the costs of postage and a relatively low return of completed questionnaires are known negatives. The distribution of questionnaires by means of e-mail at least partially eliminates these barriers.

At our clinic, we are currently testing the possibilities of obtaining data via a web form. A link with a unique identifier is sent to the patients by email, which allows backpairing the received answers with individual test subjects.

In 1993, Bradley evaluated that most of the more widely used methods of measuring results in hand surgery are rarely well validated and not necessarily well used [12]. In the last 20 years, the validity of the instruments commonly used has been gradually verified. However, it is not always possible to verify a method of measurement, because
in some cases the methodology of performing some of these tests is not sufficiently precise. Moberg has described problems with measurement accuracy in sensory tests with the words "press until the boss is happy", meaning that it is possible to achieve the desired result if the test protocol is not described precisely enough [13].

PROMs

PROMs, or Patient-reported Outcome Measures, are questionnaires that collect information about health status and treatment outcomes directly from patients. Questionnaires have been appearing in literature since the 1960s, but the term PROM does not appear until the turn of the millennium. Initially, they were developed for research purposes, especially for clinical testing of treatment effectiveness and mapping of mental health-related conditions. A review of the recent literature reveals more than 450 PROMs, of which the largest percentage are generic questionnaires assessing general quality of life, as well as questionnaires assessing musculoskeletal disorders and questionnaires specific to cancer patients. In 2019, Churruca et al. identified 42 PROMs related to musculoskeletal problems with an emphasis on only 3 groups of diseases, which were "back pain" and related problems, osteoarthritis and rheumatoid arthritis [4]. Questionnaires developed for children and adolescents constitute another chapter [14]. In the last few decades, the patient's view has been increasingly perceived as an essential aspect of evaluating the quality of health care provided. This trend has had a positive impact on the number of PROMs used as well as on the development of new ones. The high number of citations of original validation studies indicates the widespread use of PROMs. At the same time, it is possible to find many articles comparing individual PROMs among each other [3,4,15–17].

PROMs are standardized questionnaires that collect information about health status and health care outcomes, including questions about symptoms, health-related quality of life, and functional status [18]. The success of the use of PROMs depends on consistent development, patient relevance and good validation [19]. In recent years, voices have arisen, calling for a change in the perception of the concept of validation, which should become a continuous process of gathering evidence responding to ongoing socio-cultural changes, rather than a simple statement that a given questionnaire is valid [20]. When choosing a PROM, it should be borne in mind that, despite their undeniable qualities, questionnaires are only an additional source of information to support the power of objective clinical data. The choice of questionnaire to be used must depend on the desired purpose. For example, some questionnaires designed for patients with rheumatoid arthritis focus on quality-of-life assessment, others on the impact of the symptoms and others on functional status and the presence of specific symptoms [4].

Below are some of the most used questionnaires (in hand surgery), along with a brief description of the possibilities of their use.

Patient Evaluation Measure (PEM)

Developed by the British Society of Hand Surgery in the 1990s, it consists of 11 questions, with a maximum of 77 points. Less focused on function than DASH (see below), it 5 items relate to the following qualities: manual skills, movement, strength, daily activities and work; 3 evaluate symptoms (sensitivity to cold, pain and touch) and the remaining 2 are targeted at the appearance and general condition of the hand. It is proving to be one of the best tools for evaluating the results of hand disability therapy [3].

Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH)

A questionnaire developed in the mid-1990s. It assesses upper limb symptoms (e.g. pain, numbness and tingling), as well as physical, social and psychological disabilities. Optional modules related to work and sport are available. It mainly assesses function (23 out of 30 questions concern a certain manual activity, 4 questions deal with symptoms and 3 with an overall feeling of the limb). Its validity, reliability and sensitivity have been repeatedly verified in the literature. The main advantage is the official translation into many languages, incl. Czech [21,22].

QuickDASH

An abbreviated form of the DASH questionnaire with a comparable validity [21,23]. Like the original DASH, this questionnaire also exists in many official language versions, incl. Czech.

Michigan Hand Questionnaire

This questionnaire, developed in the late 1990s, evaluates overall hand function, ADL, pain, work performance, aesthetic aspects and satisfaction with hand function in 37 questions, divided into 6 modules. It can be used to assess results after surgery. In several items, it deals with hand dominance. Also available in an abbreviated form (Brief Michigan Hand Questionnaire – BMHQ) [21,24].

Boston Carpal Tunnel Questionnaire (BCTQ)

A questionnaire developed in the early 1990s in the USA. A total of 19 questions divided into 2 blocks assess the severity of subjective difficulties and the functional state of the hand [25]. It is possible to obtain a Czech translation via the website of the Czech Society of Hand Surgery, but its validity has not yet been verified in the literature [26].

Hand Unit ADL Questionnaire

Covers the function of the upper and lower limbs, together with questions of general nature. For more detailed work with the patient, it is appropriate to
Disease Activity Score in 28 Joints (DAS28)
A questionnaire specifically evaluating disease activity in rheumatoid arthritis. There are online calculators available. The number of swollen and painful joints (0–28 swollen or painful joints), the level of sedimentation or the level of C-reactive protein (CRP) and the overall health of the patient are evaluated. DAS28 is one of the core parameters for determining the severity of the disease and the starting point for choosing a treatment strategy. Theoretically, the lowest and highest scores are 0 and 10, respectively. With a DAS28 score < 2.6, we talk about the state of remission. When the DAS28 value is > 2.6, we perceive the disease as active and further divide it into 3 degrees: low (2.6–3.2), moderate (3.2–5.1) and high (> 5.1) activities [30].

Fries Health Assessment Questionnaire
A questionnaire developed in the 1980s which evaluates the functional aspects of everyday life in 8 modules (getting dressed and grooming, getting up, eating, walking, hygiene, reach, grip and activities such as housework, shopping and travelling). For each category, there are four possible answers: without any difficulty, with some difficulty, with much difficulty and unable to do. The presence of pain is automatically scored as the highest degree of disability. Fries’s assessment is considered by its proponents to be as sensitive in assessing deterioration as the changes seen in X-rays. In systemic disease, the score worsens by 0.1/year [12,31].

36-Item Short Form Health Survey (SF-36)
A universal questionnaire developed in 1992, which in eight sections deals with the overall quality of life of test subjects [32]. The biggest weakness is the effort to evaluate the sum of variously weighted areas of interest in the form of a single number [2]. There is an official Czech translation from the initiative of the Institute of Health Information and Statistics, but we were not able to verify the information on its validation [33]. It is relatively unsuitable for use in hand surgery because of its generality.

Complicating factors
Whether we use objective measurements or PROMs to evaluate health care outcomes, it is always necessary to consider the limitations of the metrics used. In both cases, the achieved results can be influenced by motivation, either on the side of the patient or on the side of the examiner. For various reasons, patients may want to achieve a worse or better score, whether with a view to obtaining insurance settlement or other compensation, or vice versa, when trying to return to a workplace where strict health regulations apply. Likewise, the examiner may be motivated to evaluate the result of care as a better or, conversely, worse score (e.g. in the case of borderline indications for surgery). To avoid these potential sources of bias, one must strictly adhere to the methodology of individual tests, and, in case of doubt, repeat testing or use other measurement methods.

Another relatively difficult outcome to compare is the cosmetic aspect. Not only the differences between sexes, but even the perception of certain disabilities and their consequences based on ethnic differences must be considered. Finally, the topic of complicating factors of measurement accuracy raises a debate on the topic: “What exactly is patient satisfaction?” Measuring patient satisfaction is certainly a useful indicator of the overall quality of care [2]. Nevertheless, a study regarding patient satisfaction by Williams and Cullan shows that although 95% of respondents say they were satisfied with their GP’s care, as many as 38% felt they could not adequately discuss their personal problems with them, 26% were dissatisfied with the information they
were provided, and a full quarter of respondents were dissatisfied with the duration of the consultation [34]. It is therefore questionable what satisfaction is from the patient’s point of view and what it is from the evaluator’s point of view.

**Conclusion**

Evaluation of results is an integral part of providing health care, both acute and elective. Only an adequately chosen method can provide relevant results, which can then be used to compare the quality of care across literature and can possibly lead to the adjustment of treatment procedures. Several more or less objective tests can be used to evaluate the results in hand surgery [35]. Since about the 1980s, the evaluation of patients’ subjective perceptions has been applied more frequently. It is possible to use several questionnaires, so-called PROMs, which are repeatedly validated. Especially with regard to the COVID-19 pandemic and rising fuel prices, they provide valuable opportunities for health care assessment without the need for the patient to be physically present in the medical facility. In order to obtain relevant data, it is particularly important to select an adequate questionnaire with a sufficiently high sensitivity for the evaluated quality, verified validity and good reliability.

For the needs of hand surgery, DASH and QuickDASH seem to be the most versatile PROMs. At the same time, they both have the advantage in the availability of an official Czech translation. PR(W)HE would also be practical for introduction into Czech practice, providing the possibility of a good comparison with the rest of the world due to its frequent use. Finally, it would be appropriate, with regard to the frequency of the diagnoses they were developed for, to translate and validate the FIHOA and BCTQ questionnaires for the evaluation of problems based on osteoarthritis and carpal tunnel syndrome.

**Disclosure:** There are no conflicts of interest in relation with the theme, creating and publication of this manuscript. The preparation of this manuscript was in accordance with the principles of the Declaration of Helsinki.

**Funding:** The author has no financial disclosures to declare.

**Role of author:** The author has written the text on his own.

**References**

23. Gummesson C., Ward MM., Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. BMC Musculoskelet Disord. 2006, 18(7): 44.
26. Dotazník BCTSQ. Praha: Česká společnost chirurgie ruky ČLS JEP; c2009 [online]. Available from:
https://d6scj24zvbbo.cloudfront.net/5193a-02fb3ef0701998b347dd209374b/200000145-3c3a53d382/Boston%20Carpal%20tunnel%20syndrome%20questionnaire.pdf.


