The Significance of Satisfaction with Treatment Results

A Data-Driven Approach Towards Personalized Healthcare

Willemijn Anna de Ridder

Over the course of this thesis (financial) support was received from:
Xpert Clinics Hand and Wrist Care, Xpert Hand Therapy, Dept. of Rehabilitation Medicine and Dept. of Plastic Surgery Erasmus MC Rotterdam, Erasmus University, WeDesign, Van Wijngaarden Medical, Chipsoft, BAP Medical, Scientific College Physical Therapy (WCF) of the Royal Dutch Society for Physical Therapy (KNGF), Tromp Medical BV, and the Dutch Society for Hand Therapy (NVHT).

Printing:	Ridderprint
Layout and design:	Wiebke Keck, persoonlijkproefschrift.nl
Cover Design:	Harma Makken, persoonlijkproefschrift.nl
ISBN:	978-94-6506-461-1

The Significance of Satisfaction with Treatment Results
A data-driven approach towards personalized healthcare

De significantie van tevredenheid met het behandelresultaat
Een datagedreven benadering voor gepersonaliseerde zorg

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. A.L. Bredenoord

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

Woensdag 20 november 2024 om 15:30 uur

door

Willemijn Anna de Ridder
geboren op 6 december 1984 te Gaasterland

Erasmus University Rotterdam

Promotiecommissie

Promotor:	Prof. dr. R.W. Selles
Overige leden:	Prof. dr. J.J. van Busschbach
	Prof. dr. ir. M.S. Kleinsmann
	Prof. dr. C. Veenhof
Copromotor(en):	Dr. H.P. Slijper
	Dr. G.M. Vermeulen
Paranimfen:	Drs. Y.E. van Kooij
	Drs. H. Koster

CONTENT

1

GENERAL INTRODUCTION AND THESIS OUTLINE

General introduction and thesis outline

“Are you satisfied?” You have likely been asked to rate your satisfaction with a product or service multiple times - whether it be after using a restroom at the airport, buying a bag online, or requesting help with a technological problem. The smiley faces may become repetitive and annoying, but you answer anyway in the hope that someone will act on your feedback and improve the quality of the product or service. However, the reality of satisfaction assessment is much more complex than simply choosing between a happy or sad face. What exactly are you satisfied with? The clean restroom or the fact that you could use a changing table for your child? The quality of the product you ordered or the time to delivery? The answer to your technical problem or the friendly approach of the employee? And what if other factors in the restaurant, like a boring date, the death of your pet, or too loud music, impact your ability to judge the food you’re eating? What if you’ve never tried sushi before - would you still be able to provide an accurate assessment of the food? And what if the same food is served in a star restaurant or in a sports canteen – would expectations influence your evaluation? All these questions highlight the complexity of rating and interpreting satisfaction, and the potential influence of factors such as context, mental health, expectations, and life events. Also, these questions immediately raise the issue of how you can properly measure satisfaction. Satisfaction is a continuum, in which a person can also be a little satisfied. It is clear that satisfaction cannot be captured in just a happy or sad smiley, but how can we measure it, then?

This thesis focuses on satisfaction with the results of treatment in a medical setting, specifically in patients with hand or wrist conditions. Understanding satisfaction with treatment results (STR) is an increasingly relevant topic in the shift toward patient-centered and value-based care. These frameworks focus, amongst others, on outcomes relevant to the patient and on achieving better outcomes at lower costs[1-5]. STR is a crucial aspect of patient-centered care; it is essential to understand patients’ perceptions of their results and to identify areas for improvement of care.

When evaluating satisfaction, it is important to distinguish between different types of satisfaction, such as satisfaction with the process of care and satisfaction with the achieved results of treatment. These different kinds of satisfaction are related but separate concepts. The first can be compared to your experience of a service delivered, such as the way the waiter served your food. The second, however, relates to the actual food served. In healthcare, for example, a patient may be satisfied with the results of their treatment but unsatisfied with the pre-treatment information provision. Conversely, a patient may consider the results of their treatment unsatisfactory but be satisfied with the empathy and clear communication of their clinician.

STR is an important outcome for patients, regardless of their condition or treatment. Many treatments for patients with hand and wrist conditions are elective and intended to improve patient-reported outcomes such as pain, ability to perform activities of daily life, or quality of life[6,7]. In this case, patients have several treatment options, including no treatment. This is in contrast to life-saving medical procedures for mortal illness or after trauma, for example. In elective and preference-sensitive treatments, where patients are free to choose and there is no necessity for treatment, STR may be even more important. After all, why choose a technically perfect procedure with a negligible risk of complications and excellent objective outcomes if you're not going to be satisfied with the result?

The aim of this thesis was to enhance patient-centered and value-based care by improving satisfaction with treatment results for patients with hand or wrist disorders. To achieve this, we aimed to:

1. Develop a more comprehensive understanding of satisfaction with treatment results and its related factors in patients with hand or wrist disorders
2. Explore the connection with the patient's mindset
3. Improve satisfaction with treatment results using data-driven tools

To meet these aims, the thesis is structured into three parts.

Part 1: Measure and understand satisfaction with treatment results

Measuring STR is a crucial aspect of evaluating the effectiveness of medical treatments. Before the shift to patient-centered care, the success of treatment was generally evaluated with clinician-reported outcomes, such as grip strength, range of motion, nerve conduction measurements, or radiological findings. We now know that clinician-reported outcome measurements (CROMs) do not necessarily correlate strongly with patient-reported outcomes measurements (PROMs) or with the patient's perception of a successful treatment[8,9]. This means that even if a clinician reports good results based on their own measurements, the patient may not necessarily feel satisfied or consider their treatment successful. The discrepancy between CROMs and PROMs emphasizes the need to use PROMs, including one on STR, to capture the patient's experience and perception of treatment success[10,11].

Worldwide, variations on the question "How satisfied are you with the results of the treatment?" are used to measure STR. The answer options may be dichotomous (yes/no), categorical (e.g., using a 5 or 7-point Likert scale), or continuous (e.g., using a Visual Analogue Scale). Additionally, related questions such as "Would you be willing to undergo

the treatment again?" and "Would you recommend this treatment to your friends and family?" may be used. However, the psychometric quality of questionnaires has not improved over the last two decades, according to a recent literature review[12], and the validity and reliability of many of these questionnaires have rarely been investigated.

It is a challenge to understand STR because of its multi-dimensional nature and the various factors influencing it. A useful model for understanding these factors is the World Health Organization-supported adaptation of the International Classification of Functioning, Disability and Health (ICF)(Figure 1)[13]. Many studies have provided evidence for this biopsychosocial model[8,13-32]. In a literature review, Marks et al. found associations of STR with pain and symptoms, activities of daily living or function, esthetics, embodiment, strength, ROM, fulfillment of expectations, deformity, workers' compensation status, and length of follow-up[13].

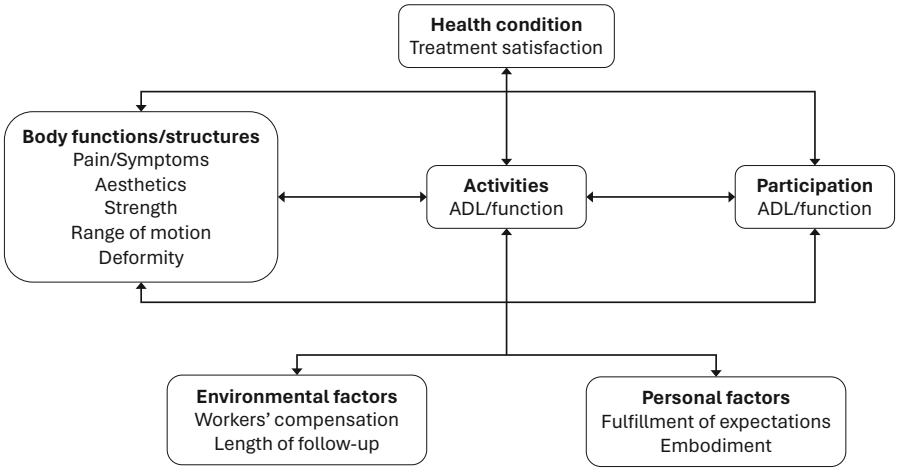


Fig. 1. Representation of how the concept of satisfaction with treatment results (STR, top box) can be integrated within the ICF model, where at least mild correlations between the factor (all boxes below) and STR could be demonstrated. The figure is a World Health Organization-supported adaptation of the standard ICF classification. ADL = Activities of Daily Living[13].

The ICF model views human functioning comprehensively across body functions and structures, activities, and participation domains and assumes that these are influenced by environmental and personal factors and vice versa. STR is a comprehensive concept in the ICF model, encompassing a general outcome across all domains. The ICF encompasses a wide array of categories to characterize body functions, structures, activities, and participation. Additionally, environmental factors are classified as they can act as either barriers or facilitators to functioning. Contrarily, personal factors, e.g., comprising coping mechanisms, education, and behavioral patterns, remain unclassified due to large social and cultural differences[33]. Patient experiences may not have been considered for the same reason, but research has found associations of STR with both. For example, strong

links have been established between STR and improved patient-reported experience measures (PREMs), such as provision of general information, treatment information, and shared decision-making[8,14-18]. Additionally, the relationship with the surgeon, particularly the perceived empathy, is a major determinant of STR[19-22]. There are also several associations between satisfaction and mental health, e.g., depressed patients rapport lower satisfaction[23-25]. Finally, there is growing evidence for the association between more positive outcome expectations and higher STR[15,29-32]. Although many associations of all these different factors with STR have been explored, there was a lack of comprehensive understanding of these relationships due to small study samples and univariable analysis[19,25,34]. Thus, studies with large sample sizes conducting multivariable analyses are needed to obtain reliable estimates for factors that are independently associated with STR.

This leads to the aims of Part 1:

- To investigate the psychometric properties of measures for evaluating satisfaction with treatment results.
- To identify factors associated with satisfaction with treatment results.

Part 2: Explore the connection with the patient's mindset

As previously stated, the patient's mindset is associated with STR. The mindset can be defined as the set of attitudes held by someone, where attitudes include a way of thinking or feeling about someone or something reflected in a person's behavior[35]. One aspect of the mindset is the expectations the patient holds towards the outcomes of a treatment. Recent studies have shown that positive expectations are associated with better outcomes and higher satisfaction, e.g., in patients undergoing hand or wrist treatment[30,36]. Outcome expectations are believed to play a crucial role in the placebo effect, which refers to the non-specific therapeutic effects of a treatment that arise from the overall therapeutic context, including patient- and clinician-specific factors, and the interaction between the patient, clinician, treatment setting, and treatment[37-39]. For example, the surgeon's white coat adds to the contextual nonspecific effect and, on average, improves positive outcome expectations. Since higher expectations are associated with better outcomes and higher satisfaction, developing interventions to change the way care is delivered may boost expectations and thereby improve STR.

Another aspect of the mindset is mental health. In this thesis, I assessed three parameters of mental health, all associated with STR: psychological distress, pain catastrophizing, and illness perception.

- Psychological distress refers to a negative emotional state that is characterized by feelings of anxiety, depression, and stress[40].

- Pain catastrophizing is a psychological construct that refers to the negative thoughts and feelings that individuals experience when in pain[41].
- Illness perception refers to the thoughts, beliefs, and attitudes that a person holds about their health condition, including the nature of the illness, causes, symptoms, the timeline of the illness, personal control over the illness, and its impact on their life[35].

Theoretically, improving patient's mindset should lead to improved STR and other outcomes. As clinicians play a vital role in communicating and addressing patient concerns during their consultation, there may be opportunities to influence certain aspects of mental health and thereby indirectly improve STR. However, this relationship had yet to be fully investigated.

This leads to the aims of Part 2:
- To identify factors associated with pre-treatment outcome expectations
- To evaluate the change in mental health following the first hand surgeon consultation

Part 3: Improve Satisfaction with Treatment Results using Data-Driven Tools

Although STR is a crucial aspect of patient-centered care and considered an essential outcome domain in patients with hand or wrist conditions[42], it is important to realize that it should not be a goal in itself. For example, in many situations, healthcare providers should strive to provide patients with the best long-term solution, even if this doesn't lead to high levels of satisfaction in the short term. Think of surgery to prevent future worsening or return of symptoms or of situations where the surgeon advises hand therapy or no treatment, despite the patient's wish for surgery. When designing interventions to enhance STR, we developed a schematic outline (Figure 2) to help us identify starting points for these interventions.

The schematic outline distinguishes between factors influencing STR that cannot be modified and factors that can be modified. Nonmodifiable factors include sociodemographics, such as age or type of work, and medical history and diagnosis. Modifiable factors include PROMs, PREMs, expectations, and mental health. Interventions can focus on the four modifiable factors, e.g., the way information is provided or the clinician's response to depression. Intervention refers to any action aimed at improving STR.

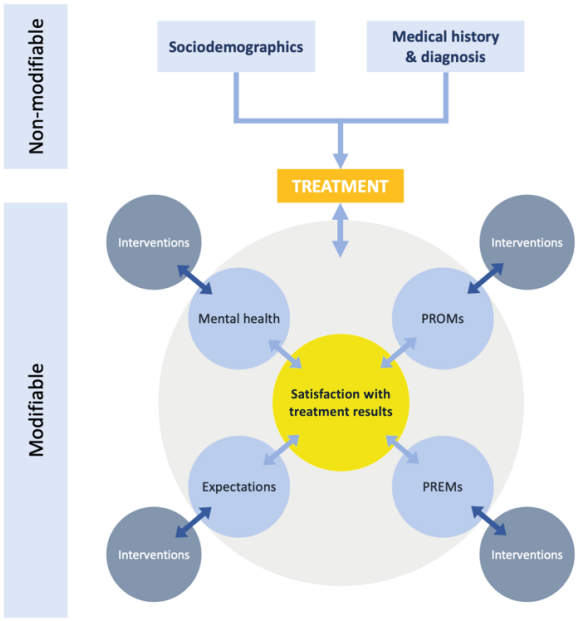


Fig. 2. Identification of starting points for interventions to improve patient-centered healthcare with satisfaction with treatment results (STR). Nonmodifiable factors include sociodemographics, medical history, and diagnosis. Modifiable factors include PROMs, PREMs, expectations, and mental health. Interventions can focus on modifiable factors. For example, an intervention can be designed to improve the experience of the patients with the healthcare delivery (PREMS), and the improved experience due to this intervention may improve the STR. Similarly, an intervention may decrease patients' distress (Mental Health), thereby improving satisfaction. Any action aimed at improving STR is considered an intervention.

In our pursuit of developing interventions that enhance STR, we recognize that treatment selection is only partially modifiable and heavily influenced by factors such as biomechanical, personal, work-related, or financial reasons. To ensure that treatment plans are tailored to the individual patient's needs and preferences, shared decision-making between the clinician and patient is crucial. Interestingly, data-driven tools that support shared decision-making can provide a more evidence-based approach to improving healthcare. By leveraging advanced analytics, clinicians can gain a more comprehensive understanding of patient needs, leading to more informed and effective treatment decisions. This approach may help minimize unnecessary procedures and optimize resource utilization[43,44]. Therefore, in this thesis, we developed, implemented, and evaluated two data-driven tools: the Utra-Short Mental Health Screening Tool and the Patient-Specific Needs Evaluation (PSN).

To understand the background of the mental health screener, it is important to realize that the relevance of mental health in musculoskeletal healthcare has been demonstrated in

numerous studies[45-62], and it is valuable to routinely examine a patient's mental health to support personalized and value-based healthcare. However, implementing available patient-reported measures of mental health in clinical practice can be challenging due to time constraints and patient burden. Patients may not understand why they have to complete these questionnaires if, in their opinion, they have very objectifiable symptoms because of a specific physical condition (such as osteoarthritis). Consequently, patients may feel that using elaborate measures to evaluate mental health is inappropriate. Therefore, there was a need for a short screening tool that provides an accurate view of patients' mental health with a low patient and clinician burden to overcome these issues.

To understand the background of the PSN, it is vital to understand that one of the challenges in patient-centered healthcare is understanding and addressing each patient's unique needs and goals. While clinicians strive to provide the best care possible, there may be a discrepancy between what the patient needs or wants and what the clinician is able to deliver or sees as the clinical priority. This discrepancy can result in a treatment plan that does not fully align with the patient's needs or goals. To address this gap, we developed a brief patient-reported tool to assess patient-specific information needs, treatment goals, and Personal Meaningful Gain (PMG) before a first clinician consultation: the Patient-Specific Needs Evaluation (PSN).

The provision of targeted information improves shared-decision making and all sorts of outcomes, including STR[63]. Providing information and fulfilling information needs is particularly important in elective treatments, as the decision to treat often depends on this information. However, before this thesis, we were not aware of a PROM to measure these individual information needs, and the determinants of satisfactory information provision or fulfillment of information needs were not yet fully understood. Furthermore, although there exist many goalsetting instruments, none of them briefly measures the improvement a patient wants to obtain to be satisfied with the treatment result on a domain considered as most important by the patient.

To further engage in shared decision-making successfully and to tailor the treatment plan to the individual patient, clinicians need to be aware of their patient's goals and preferred outcomes. In this process, it's important to keep in mind that not all statistically significant changes in PROMs may be important to the individual patient. This is where the concept of the Minimally Important Change (MIC) and Patient Acceptable Symptom State (PASS) come into play, as they aim to address the clinical relevance of outcomes[64,65]. The MIC refers to the smallest change from the beginning of the treatment to a certain time point post-treatment that patients (on average) perceive as important. The PASS refers to a certain end state that patients (again, on average) perceive as satisfactory. However, simply reaching the MIC or PASS may not necessarily lead to STR because this may not have been what the individual patient needed. To truly put the patient at the center of healthcare, a construct was needed that evaluates the minimal improvement

that is meaningful to the individual based on a domain chosen as most important by that individual. Therefore, we developed the concept of Personal Meaningful Gain (PMG). The PMG represents the improvement an individual patient wants to obtain on a domain chosen as most important to be satisfied with the treatment results, given the baseline score. In this way, shared decision-making can be informed by an individualized understanding of what constitutes a meaningful improvement, leading to better decision-making and, ultimately, improved STR.

This leads to the aim of Part 3:

- To develop and evaluate tools that help clinicians during daily clinical care to positively respond to each individual patient's mental health, personal information needs, treatment goals, and desired improvements to improve satisfaction with treatment results

Aim and structure of this thesis

In summary, the aim of this thesis was to enhance patient-centered and value-based care by improving satisfaction with treatment results for patients with hand or wrist disorders. Aligned with the principles of these frameworks, the overarching goal is to enhance patients' well-being while optimizing the balance between improved outcomes and cost-effectiveness. This thesis underscores the importance of prioritizing patients' unique needs, values, and goals in all medical decisions and interventions.

To achieve this, we aimed to:

1. Develop a more comprehensive understanding of satisfaction with treatment results and its related factors in patients with hand or wrist disorders
2. Explore the connection with the patient's mindset
3. Improve satisfaction with treatment results using data-driven tools

The thesis is structured into three parts.

Part 1: Measure and Understand Satisfaction with Treatment Results

This section focuses on measuring and understanding satisfaction with treatment results. It includes an investigation of the psychometric properties of the Satisfaction with Treatment Result Questionnaire (**Chapter 2**) and an analysis of the factors associated with satisfaction (**Chapter 3**).

Part 2: Explore the Connection with the Patient's Mindset

Part Two examines two crucial themes linked to satisfaction with treatment results, namely patient expectations and mental health. **Chapter 4** focuses on identifying factors associated with patient pre-treatment outcome expectations, while **Chapter 5** explores the impact of the first surgeon consultation on the patient's mental health.

Part 3: Improve Satisfaction with Treatment Results using Data-Driven Tools

Part Three introduces data-driven tools to improve patient satisfaction with treatment results. **Chapter 6** describes the development of an ultra-short mental health screener, which can be used for improved shared decision-making. **Chapter 7** details the development and validation of the Patient-Specific Needs Evaluation, which identifies the patient's information needs, treatment goals, and improvement goals. **Chapter 8** analyzes the factors affecting the provision and fulfillment of the patient's information needs. **Chapter 9** explores the Personal Meaningful Gain, which is the minimum improvement needed to satisfy the patient with the treatment results, and identifies factors explaining the Personal Meaningful Gain value. In **Chapter 10**, the Personal Meaningful Gain is compared to the MIC and the PASS.

References

1. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477-2481.
2. Barry MJ, Edgman-Levitan S. Shared Decision Making — The Pinnacle of Patient-Centered Care. *New Engl J Med.* 2012;366(9):780-781.
3. Basch E. Patient-Reported Outcomes — Harnessing Patients' Voices to Improve Clinical Care. *New Engl J Med.* 2017;376(2):105-108.
4. Bernstein DN, Calfee RP, Hammert WC, Rozental TD, Witkowski ML, Porter ME. Value-Based Health Care in Hand Surgery: Where Are We & Where Do We Go From Here? *The Journal of Hand Surgery.* 2022.
5. Institute of Medicine Committee on Quality of Health Care in A. In: *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington (DC): National Academies Press (US) Copyright 2001 by the National Academy of Sciences. All rights reserved.; 2001.
6. Kamal RN. Quality and Value in an Evolving Health Care Landscape. *J Hand Surg Am.* 2016;41(7):794-799.
7. Kamal RN, Kakar S, Ruch D, Richard MJ, Akelman E, Got C, Blazar P, Ladd A, Yao J, Ring D. Quality measurement: A primer for hand surgeons. *J Hand Surg (USA).* 2016;41(5):645-651.
8. Mandl LA, Galvin DH, Bosch JP, George CC, Simmons BP, Axt TS, Fossel AH, Katz JN. Metacarpophalangeal arthroplasty in rheumatoid arthritis: What determines satisfaction with surgery? *J Rheumatol.* 2002;29(12):2488-2491.
9. Karnezis IA, Fragkiadakis EG. Association between objective clinical variables and patient-rated disability of the wrist. *J Bone Joint Surg Br.* 2002;84(7):967-970.
10. Goldhahn J, Angst F, Simmen BR. What counts: outcome assessment after distal radius fractures in aged patients. *J Orthop Trauma.* 2008;22(8 Suppl):S126-130.
11. van Oosterom FJ, Ettema AM, Mulder PG, Hovius SE. Impairment and disability after severe hand injuries with multiple phalangeal fractures. *J Hand Surg Am.* 2007;32(1):91-95.
12. Wilson CD, Mand D, Ring D, Ramtin S. A Systematic Review of Satisfaction Measures in Hand and Wrist Surgery. *J Hand Surg Am.* 2023;48(1):1-8.
13. Marks M, Herren DB, Vliet Vlieland TPM, Simmen BR, Angst F, Goldhahn J. Determinants of patient satisfaction after orthopedic interventions to the hand: A review of the literature. *J Hand Ther.* 2011;24(4):303-312.
14. Tsehaie J, van der Oest MJW, Poelstra R, Selles RW, Feitz R, Slijper HP, Hovius SER, Porsius JT. Positive experience with treatment is associated with better surgical outcome in trapeziometacarpal osteoarthritis. *J Hand Surg Eur Vol.* 2019;44(7):714-721.
15. Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet.* 2001;357(9258):757-762.
16. Parrish RC, Menendez ME, Mudgal CS, Jupiter JB, Chen NC, Ring D. Patient Satisfaction and its Relation to Perceived Visit Duration with a Hand Surgeon. *J Hand Surg (USA).* 2016;41(2):257-262e254.
17. Bot AG, Bossen JK, Herndon JH, Ruchelsman DE, Ring D, Vranceanu AM. Informed shared decision-making and patient satisfaction. *Psychosomatics.* 2014;55(6):586-594.

18. Bernstein J, Kupperman E, Kandel LA, Ahn J. Shared Decision Making, Fast and Slow: Implications for Informed Consent, Resource Utilization, and Patient Satisfaction in Orthopaedic Surgery. *J Am Acad Orthop Surg.* 2016;24(7):495-502.

19. Pollak KI, Alexander SC, Tulsky JA, Lyna P, Coffman CJ, Dolor RJ, Gulbrandsen P, Ostbye T. Physician empathy and listening: associations with patient satisfaction and autonomy. *J Am Board Fam Med.* 2011;24(6):665-672.

20. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Health Prof.* 2004;27(3):237-251.

21. Menendez ME, Chen NC, Mudgal CS, Jupiter JB, Ring D. Physician Empathy as a Driver of Hand Surgery Patient Satisfaction. *J Hand Surg (USA).* 2015;40(9):1860-1865.

22. Vranceanu AM, Ring D. Factors associated with patient satisfaction. *J Hand Surg (USA).* 2011;36(9):1504-1508.

23. Mosegaard SB, Stilling M, Hansen TB. Higher preoperative pain catastrophizing increases the risk of low patient reported satisfaction after carpal tunnel release: a prospective study. *BMC Musculoskelet Disord.* 2020;21(1):42.

24. Mosegaard SB, Stilling M, Hansen TB. Pain Catastrophizing Scale as a predictor of low postoperative satisfaction after hand surgery. *J Orthop.* 2020;21:245-248.

25. Lozano Calderón SA, Paiva A, Ring D. Patient Satisfaction After Open Carpal Tunnel Release Correlates With Depression. *J Hand Surg (USA).* 2008;33(3):303-307.

26. Marks M, Audigé L, Reissner L, Herren DB, Schindele S, Vliet Vlieland TPM. Determinants of patient satisfaction after surgery or corticosteroid injection for trapeziometacarpal osteoarthritis: Results of a prospective cohort study. *Arch Orthop Trauma Surg.* 2015;135(1):141-147.

27. Kadzielski J, Malhotra LR, Zurakowski D, Lee SGP, Jupiter JB, Ring D. Evaluation of Preoperative Expectations and Patient Satisfaction After Carpal Tunnel Release. *J Hand Surg (USA).* 2008;33(10):1783-1788.

28. Frouzakis R, Herren DB, Marks M. Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis. *J Hand Surg [Am].* 2015;40(3):483-490.

29. Hoogendam L, van der Oest MJW, Wouters RM, Andrinopoulou ER, Vermeulen GM, Slijper HP, Porsius JT, Selles RW. Patients with higher treatment outcome expectations are more satisfied with the results of nonoperative treatment for thumb base osteoarthritis: a cohort study. *Arch Phys Med Rehabil.* 2021;102(8):1533-1540.

30. Mandl LA, Burke FD, Shaw Wilgis EF, Lyman S, Katz JN, Chung KC. Could preoperative preferences and expectations influence surgical decision making? Rheumatoid arthritis patients contemplating metacarpophalangeal joint arthroplasty. *Plast Reconstr Surg.* 2008;121(1):175-180.

31. Swarup I, Henn CM, Gulotta LV, Henn RF, 3rd. Patient expectations and satisfaction in orthopaedic surgery: A review of the literature. *J Clin Orthop Trauma.* 2019;10(4):755-760.

32. Poelstra R, Selles RW, Slijper HP, van der Oest MJW, Feitz R, Hovius SER, Porsius JT. Better patients' treatment experiences are associated with better postoperative results in Dupuytren's disease. *J Hand Surg Eur Vol.* 2018;43(8):848-854.

33. Organization WH. International Classification of Functioning, Disability and Health (ICF). https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health. Published 2001. Accessed.

34. Ring D, Kadzielski J, Fabian L, Zurakowski D, Malhotra LR, Jupiter JB. Self-reported upper extremity health status correlates with depression. *J Bone Joint Surg Am.* 2006;88(9):1983-1988.

35. OED OED. Oxford english dictionary. *Online edition at< dictionary oed com/entrance dtl.* 2010.

36. van der Oest MJW, Hoogendam L, Wouters RM, Vermeulen GM, Slijper HP, Selles RW, Vranceanu AM, Porsius JT, Hand-Wrist Study G. Associations between positive treatment outcome expectations, illness understanding, and outcomes: a cohort study on non-operative treatment of first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021:1-8.

37. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol.* 2008;59:565-590.

38. Miller FG, Kaptchuk TJ. The power of context: reconceptualizing the placebo effect. *J R Soc Med.* 2008;101(5):222-225.

39. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Biological, clinical, and ethical advances of placebo effects. *Lancet.* 2010;375(9715):686-695.

40. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* 2009;50(6):613-621.

41. Sullivan MJL BS, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychol Assess.* 1995;7:524-521.

42. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, Hovius SER. A Standard Set for Outcome Measurement in Patients With Hand and Wrist Conditions: Consensus by the International Consortium for Health Outcomes Measurement Hand and Wrist Working Group. *J Hand Surg Am.* 2021;46(10):841-855.e847.

43. Belle A, Thiagarajan R, Soroushmehr SM, Navidi F, Beard DA, Najarian K. Big Data Analytics in Healthcare. *Biomed Res Int.* 2015;2015:370194.

44. Keller M, Guebeli A, Thieringer F, Honigmann P. Artificial intelligence in patient-specific hand surgery: a scoping review of literature. *Int J Comput Assist Radiol Surg.* 2023.

45. Becker SJ, Makarawung DJ, Spit SA, King JD, Ring D. Disability in patients with trapeziometacarpal joint arthrosis: incidental versus presenting diagnosis. *J Hand Surg Am.* 2014;39(10):2009-2015.e2008.

46. Blackburn J, van der Oest MJW, Selles RW, Chen NC, Feitz R, Vranceanu AM, Porsius JT. Which Psychological Variables Are Associated With Pain and Function Before Surgery for de Quervain's Tenosynovitis? A Cross-sectional Study. *Clin Orthop Relat Res.* 2019;477(12):2750-2758.

47. Blackburn J, van der Oest MJW, Chen NC, Feitz R, Duraku LS, Zuidam JM, Vranceanu AM, Selles RW. Are Patient Expectations and Illness Perception Associated with Patient-reported Outcomes from Surgical Decompression in de Quervain's Tenosynovitis? *Clin Orthop Relat Res.* 2021;479(5):1147-1155.

48. Damman W, Liu R, Kaptein AA, Evers AWM, van Middendorp H, Rosendaal FR, Kloppenburg M. Illness perceptions and their association with 2 year functional status and change in patients with hand osteoarthritis. *Rheumatology (Oxford).* 2018;Dec 1;57(12):2190-2199. doi: 10.1093/rheumatology/key231.

49. Dave AJ, Selzer F, Losina E, Klara KM, Collins JE, Usiskin I, Band P, Dalury DF, Iorio R, Kindsfater K, Katz JN. Is there an association between whole-body pain with osteoarthritis-related knee pain, pain catastrophizing, and mental health? *Clin Orthop Relat Res.* 2015;473(12):3894-3902.

50. Duivenvoorden T, Vissers MM, Verhaar JAN, Busschbach JJV, Gosens T, Bloem RM, Bierma-Zeinstra SMA, Reijman M. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study. *Osteoarthritis Cartilage.* 2013;21(12):1834-1840.

51. Foster NE, Thomas E, Hill JC, Hay EM. The relationship between patient and practitioner expectations and preferences and clinical outcomes in a trial of exercise and acupuncture for knee osteoarthritis. *Eur J Pain.* 2010;14(4):402-409.

52. Frouzakis R, Herren DB, Marks M. Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis. *J Hand Surg Am.* 2015;40(3):483-490.

53. Glette M, Landmark T, Jensen MP, Woodhouse A, Butler S, Borchgrevink PC, Stiles TC. Catastrophizing, Solicitous Responses From Significant Others, and Function in Individuals With Neuropathic Pain, Osteoarthritis, or Spinal Pain in the General Population. *J Pain.* 2018;19(9):983-995.

54. Hoogendam L, van der Oest MJW, Tsehaie J, Wouters RM, Vermeulen GM, Slijper HP, Selles RW, Porsius JT. Psychological factors are more strongly associated with pain than radiographic severity in non-invasively treated first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021;43(13):1897-1902.

55. Hoogendam L, van der Oest MJW, Wouters RM, Andrinopoulou ER, Vermeulen GM, Slijper HP, Porsius JT, Selles RW. Patients With Higher Treatment Outcome Expectations Are More Satisfied With the Results of Nonoperative Treatment for Thumb Base Osteoarthritis: A Cohort Study. *Arch Phys Med Rehabil.* 2021.

56. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative Pain Catastrophizing Predicts Pain Outcome after Knee Arthroplasty. *Clinical Orthopaedics and Related Research®.* 2010;468(3):798-806.

57. Sharma A, Kudesia P, Shi Q, Gandhi R. Anxiety and depression in patients with osteoarthritis: impact and management challenges. *Open access rheumatology : research and reviews.* 2016;8:103-113.

58. van der Oest MJW, Poelstra R, Feitz R, Vranceanu AM, Slijper HP, Selles RW, Porsius JT. Illness Perceptions of Patients With First Carpometacarpal Osteoarthritis, Carpal Tunnel Syndrome, Dupuytren Contracture, or Trigger Finger. *J Hand Surg Am.* 2020;45(5):455.e451-455.e458.

59. van der Oest MJW, Hoogendam L, Wouters RM, Vermeulen GM, Slijper HP, Selles RW, Vranceanu AM, Porsius JT. Associations between positive treatment outcome expectations, illness understanding, and outcomes: a cohort study on non-operative treatment of first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021:1-8.

60. Wood TJ, Thornley P, Petruccelli D, Kabali C, Winemaker M, de Beer J. Preoperative Predictors of Pain Catastrophizing, Anxiety, and Depression in Patients Undergoing Total Joint Arthroplasty. *J Arthroplasty.* 2016;31(12):2750-2756.

61. Wouters RM, Vranceanu AM, Slijper HP, Vermeulen GM, van der Oest MJW, Selles RW, Porsius JT. Patients With Thumb-base Osteoarthritis Scheduled for Surgery Have More Symptoms, Worse Psychological Profile, and Higher Expectations Than Nonsurgical Counterparts: A Large Cohort Analysis. *Clin Orthop Relat Res.* 2019;477(12):2735-2746.

62. Wouters RM, Porsius JT, van der Oest MJW, Slijper HP, Souer JS, Selles RW, MacDermid JC. Psychological Characteristics, Female Sex, and Opioid Use Predict Acute Postoperative Pain in Patients Surgically Treated for Thumb Base Osteoarthritis: A Cohort Study. *Plast Reconstr Surg.* 2020;146(6):1307-1316.

63. Larson CO, Nelson EC, Gustafson D, Batalden PB. The relationship between meeting patients' information needs and their satisfaction with hospital care and general health status outcomes. *Int J Qual Health Care.* 1996;8(5):447-456.

64. Terwee CB, Peipert JD, Chapman R, Lai JS, Terluin B, Cella D, Griffith P, Mokkink LB. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Qual Life Res.* 2021;30(10):2729-2754.

65. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. *Arthritis Rheum.* 2006;55(4):526-530.

PART 1

MEASURE & UNDERSTAND SATISFACTION WITH TREATMENT RESULTS

2

TEST-RETEST RELIABILITY AND CONSTRUCT VALIDITY OF THE SATISFACTION WITH TREATMENT RESULT QUESTIONNAIRE IN PATIENTS WITH HAND AND WRIST CONDITIONS: A PROSPECTIVE STUDY

de Ridder WA, van Kooij YE, Vermeulen GM, Slijper HP, Selles RW, Wouters RM, and the Hand-Wrist Study Group; 2021; Clinical Orthopaedics and Related Research

Abstract

Background

A patient's satisfaction with a treatment result is an important outcome domain as clinicians increasingly focus on patient-centered, value-based health care. However, to our knowledge, there are no validated satisfaction metrics focusing on treatment results for hand and wrist conditions.

Questions/purposes

Among patients who were treated for hand and wrist conditions, (1) what is the test-retest reliability of the Satisfaction with Treatment Result Questionnaire? (2) What is the construct validity of that outcomes tool?

Methods

This was a prospective study using two samples: a test-retest reliability sample and a construct validity sample. For the test-retest sample, data collection took place between February 2020 and May 2020, and we included 174 patients at the end of their treatment with complete baseline data that included both the primary test and the retest. Test-retest reliability was evaluated with a mean time difference of 7.2 ± 1.6 days. For the construct validity sample, data collection took place between January 2012 and May 2020. We included 3750 patients who completed the Satisfaction with Treatment Result Questionnaire, VAS, and the Net Promotor Score (NPS) at 3 months. Construct validity was evaluated using hypothesis testing, in which we correlated the patients' level of satisfaction to the willingness to undergo the treatment again, VAS scores, and the NPS. We performed additional hypothesis testing on 2306 patients who also completed the Michigan Hand outcomes Questionnaire (MHQ). Satisfaction with the treatment result was measured as the patients' level of satisfaction on a 5-point Likert scale and their willingness to undergo the treatment again under similar circumstances.

Results

We found high reliability for level of satisfaction measured on Likert scale (ICC 0.86 [95% CI 0.81 to 0.89]), and almost-perfect agreement for both level of satisfaction measured on Likert scale (weighted kappa 0.86 [95% CI 0.80 to 0.91]) and willingness to undergo the treatment again (kappa 0.81 [95% CI 0.70 to 0.92]) of the Satisfaction with Treatment Result Questionnaire. Construct validity was good to excellent as seven of the eight hypotheses were confirmed. In the confirmed hypotheses, there was a moderate-to-strong correlation with VAS pain, VAS function, NPS, MHQ pain, and MHQ general hand function (Spearman rho ranging from 0.43 to 0.67; all $p < 0.001$) and a strong to very strong correlation with VAS satisfaction and MHQ satisfaction (Spearman rho of 0.73 and 0.71; both $p < 0.001$). The rejected hypothesis indicated only a moderate correlation between level of satisfaction on a 5-point Likert scale and willingness to undergo the treatment again under similar circumstances (Spearman rho of 0.44; $p < 0.001$).

Conclusion

The Satisfaction with Treatment Result Questionnaire has good-to-excellent construct validity and very high test-retest reliability in patients with hand and wrist conditions.

Clinical Relevance

This questionnaire can be used to reliably and validly measure satisfaction with treatment result in striving for patient-centered care and value-based health care. Future research should investigate predictors of variation in satisfaction with treatment results.

Introduction

The patient-centered care and value-based healthcare frameworks have gained recognition globally in recent years [1-3]. In these frameworks, the patient is central, and the aim is to achieve high value at low cost [1-3]. Using patient-reported outcome measurements (PROMs) is an important aspect of patient-centered care and value-based health care [4-8], and the degree to which a patient is satisfied with his or her treatment result may be one of the most important and relevant PROMs to use [1-3,9,10]. Satisfaction with the treatment result is measured worldwide using variations of the question "how satisfied are you with your treatment results so far?" [11], but there are doubts about the reliability and validity of measuring satisfaction with the treatment result [12-14]. The patient's opinion about the treatment result seems too difficult to measure, depending on pre-treatment expectations, environmental factors and psychological factors, among others [12].

Although the same considerations apply to any PROM, several studies have shown that many well-designed PROMs are valid and reliable [15-17]. However, to the best of our knowledge, no studies on the reliability and validity of a PROM evaluating patient satisfaction with the treatment result are currently available. Further, as satisfaction with treatment result is important for patient-centered care and value-based healthcare and it is also considered an essential outcome domain by the recently published International Consortium for Health Outcomes Measurement standard set for hand and wrist conditions [18], it is important to further investigate the psychometric properties of measures for evaluating satisfaction with treatment results.

Therefore, we asked: Among patients who were treated for hand and wrist conditions, (1) what is the test-retest reliability of the Satisfaction with Treatment Result Questionnaire? (2) What is the construct validity of that outcomes tool?

Patients and Methods

Study Design

This was a prospective study on the test-retest reliability and construct validity of the Satisfaction with Treatment Result questionnaire, using a prospective and population-based sample of patients with hand and wrist conditions from the Hand Wrist Study Group cohort. This study was reported following the Strengthening the Reporting of Observational Studies in Epidemiology statement [19].

Setting

Data were collected at Xpert Clinic and Handtherapie Nederland, currently comprising 28 clinics for hand surgery and therapy in The Netherlands. Twenty-three surgeons certified

by the Federation of European Societies for Surgery of the Hand and more than 150 hand therapists are employed at our treatment centers.

The primary data collection on satisfaction with the treatment result was part of the usual care and occurred between January 2012 and May 2020, and additional retest data for satisfaction with the treatment result were collected between February 2020 and May 2020. Data were collected using GemsTracker electronic data capture tools (GemsTracker 2020, Erasmus MC and Equipe Zorgbedrijven). GemsTracker is a secure internet-based application for the distribution of questionnaires and forms during clinical research and quality registrations. Details on the Hand Wrist Study Group cohort have been published [20].

Participants

In this study, we used two samples: a sample to evaluate the test-retest reliability and a sample to evaluate the construct validity.

Patients were eligible for inclusion in the test-retest reliability sample if they were treated for any hand or wrist condition and completed the Satisfaction with Treatment Result Questionnaire at the final timepoint of outcome measurement as defined within the Hand Wrist Study Group cohort [20]. We chose the final timepoint because we expected little or no change in health status at that timepoint. The following timepoints were included: 3 months after minor surgery or nonsurgical treatment (for example, trigger finger release or exercise therapy), 6 months after treatment for neuropathies or Dupuytren's (such as carpal tunnel release or limited fasciectomy), and 12 months after more extensive surgery (for example, thumb carpometacarpal resection arthroplasty). Five to 7 days after completing the Satisfaction with Treatment Result Questionnaire, patients were invited to complete the questionnaire again to evaluate the test-retest reliability. The retest questionnaire was available for 6 days after patients received the invitation, creating a time interval of 5 days to 13 days as we hypothesized that the construct of satisfaction with treatment result remained stable over that time frame. The average time between the primary test and retest of the Satisfaction with Treatment Result Questionnaire was 7.2 ± 1.6 days.

Patients were eligible for inclusion in the construct validity sample if they completed the Satisfaction with Treatment Result Questionnaire, VAS for pain during physical load, VAS function, VAS satisfaction with the hand, and the Net Promotor Score (NPS) at 3 months after treatment. We used 3 months as a timepoint because the NPS was only administered at this timepoint. We additionally composed a subset of patients within this sample that also completed the Michigan Hand outcomes Questionnaire (MHQ). This was only a subset of patients as this questionnaire is not administered for every patient in our cohort. We included patients who underwent one of the following common treatments: nonsurgical treatment for carpometacarpal osteoarthritis of the thumb, surgical treatment

for carpometacarpal osteoarthritis of the thumb; nonsurgical treatment for wrist tendonitis or tenosynovitis, three-ligament tenodesis (modified Brunelli procedure), trigger finger release, proximal interphalangeal joint arthroplasty, limited fasciectomy, and carpal tunnel release. These treatments were chosen because they are the most common treatments in our eight different measurement tracks [20].

For the test-retest reliability sample, we screened 330 consecutive patients who completed the Satisfaction with Treatment Result questionnaire at the final time point. Of those, 287 had complete baseline sociodemographics and were invited to complete the retest. One hundred and thirteen patients did not respond, thus we included 174 patients in the test-retest reliability sample (Table 1).

Table 1. Baseline characteristics of the first sample for test-retest reliability (n = 174 patients)

Variable	Value
Age in years	56.4 ± 12.8
Male sex	41% (72)
Dominant side treated	53% (93)
Initial (rather than second) opinion	96% (167)
Type of work	
Unemployed	35% (61)
Light physical labor	29% (51)
Moderate physical labor	31% (53)
Heavy physical labor	5% (9)
Symptom duration in months	22.6 ± 45
Measurement track	
Thumb regular	7% (12)
Thumb extended	9% (15)
Wrist regular	18% (31)
Wrist extended	10% (18)
Finger regular	19% (33)
Finger extended	5% (8)
Dupuytren's	14% (25)
Nerve compression or decompression	18% (32)
Primary test of satisfaction with treatment result: Question 1	
Excellent	22% (38)
Good	36% (62)
Fair	21% (36)
Moderate	17% (30)
Poor	5% (8)
Primary test of satisfaction with treatment result: Question 2 = no	20% (34)

Data are presented as % (n) or mean ± SD, as appropriate.

For the construct validity sample, we screened 31,846 patients with complete sociodemographics and within that group, 17,174 patients completed the Satisfaction with Treatment Result questionnaire at 3 months. Of those, 3750 patients also completed VAS and the NPS and were included in the construct validity sample (Table 2). Within the construct validity sample, 2306 patients also completed the MHQ and were included in the subanalyses for construct validity (Fig. 1)

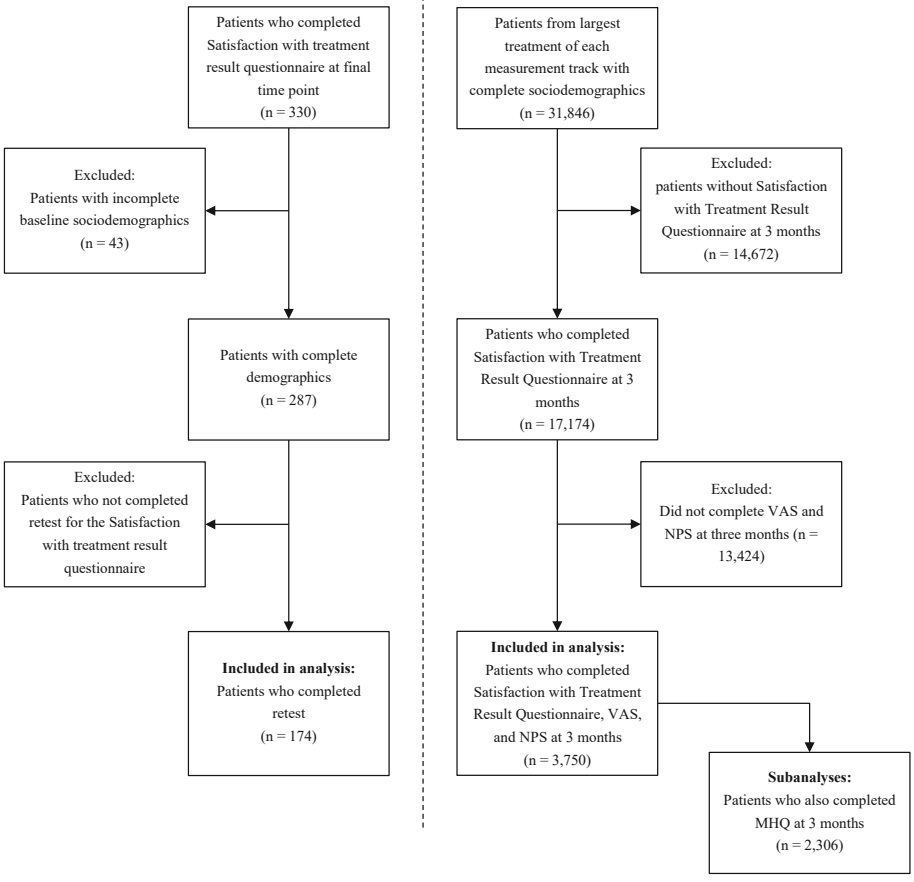


Fig. 1 This flowchart shows the patients who were included in the study. The left side displays the test-retest reliability sample, and the right side displays the construct validity sample.

Table 2. Baseline characteristics of the second sample for the entire construct validity (n = 3750) and the subset of patients that also completed the Michigan Hand outcomes Questionnaire (MHQ, n = 1692).

Variable	**Value**	
	Entire construct validity sample (n = 3750)	Subset that completed MHQ (n = 2306)
Age in years	59 ± 12	62 ± 10
Male sex	41% (1544)	46% (1067)
Dominant side treated	51% (1903)	99% (2272)
Initial (rather than second) opinion	98% (3669)	46% (1061)
Type of work		
Unemployed	41% (1554)	47% (1088)
Light physical labor	26% (979)	27% (614)
Moderate physical labor	22% (834)	19% (428)
Heavy physical labor	10% (383)	8% (176)
Treatment		
Nonsurgical treatment for CMC-1 OA	15% (573)	23% (532)
Surgical treatment for CMC-1 OA	9% (317)	13% (309)
Nonsurgical treatment for wrist tendonitis or tenosynovitis	8% (283)	-
Three-ligament tenodesis (modified Brunelli)	2% (80)	-
Trigger finger release	23% (880)	36% (830)
PIP joint arthroplasty	1% (28)	1% (21)
Limited fasciectomy	17% (627)	27% (614)
Carpal tunnel release	26% (962)	-
Symptom duration in months	21.7 ± 37	23.9 ± 38
Satisfaction with treatment result: Question 1		
Excellent	20% (766)	18% (412)
Good	38% (1434)	38% (877)
Fair	25% (940)	26% (604)
Moderate	12% (442)	14% (320)
Poor	4% (160)	4% (93)
Satisfaction with treatment result: Question 2 = no , n (%)	15% (571)	16% (379)

Data are presented as % (n) or mean ± SD, as appropriate. CMC-1 = carpometacapal of the thumb; OA = osteoarthrtitis.

Variables, Data Sources, and Measurements

The primary outcome of this study was the Satisfaction with Treatment Result questionnaire, comprising two questions. Question 1 evaluates the patient's satisfaction with the treatment result thus far, using a 5-point Likert scale (exact question: "How satisfied are you with your treatment result thus far?"; answering options were poor, moderate, fair, good, and excellent). In Question 2, the patient indicates whether he or she would undergo the same procedure again under similar circumstances (exact question: "If you would be in the same circumstances, would you be willing to undergo this treatment again?"; answering options were yes or no).

To evaluate the construct validity, we used several other questionnaires to calculate between-questionnaire correlations. We used a VAS (range 0-100), which is reliable and valid [15] to measure pain (higher scores indicate more pain), hand function (higher scores indicate better function), and satisfaction with the hand (exact question: "How satisfied are you with your hand at this moment?"; higher scores indicate greater satisfaction). We also used the MHQ subscales pain, general hand function, and satisfaction with hand (all subscales: range 0-100, higher scores indicate better performance), but as this questionnaire is not administered for every patient this was only done for a subset of the entire construct validity sample.

Finally, we used the NPS in the hypothesis testing, which is a metric to assess the quality-of-service delivery [21,22]. This included a single question indicating the extent to which patients would recommend our clinic to friends and family on a 10-point scale; higher scores indicate a stronger recommendation.

Ethical Approval

Ethical approval for this study was obtained from Erasmus MC, Rotterdam, the Netherlands (approval number 2018-1088). This study was performed in accordance with the Declaration of Helsinki and approved by the local medical research ethical committee. Written informed consent was obtained from all patients.

Study Size

A priori power analysis for the test-retest reliability sample, testing the null hypothesis (Cohen's kappa = 0.7, indicating substantial agreement) versus the alternative hypothesis (Cohen's kappa > 0.7, given that kappa = 0.85 and ratings classify 50% in agreement), suggested that a sample of 96 participants was required, which was below the included sample of 174 patients we included in the test-retest reliability sample.

For the construct validity sample, a post-hoc power analysis for the Spearman correlation, with an $\alpha = 0.05$, $\beta = 0.10$, and an expected correlation coefficient of $r = 0.20$, suggested that a sample of 259 participants was required, which was well below the included samples of 3750 patients we included in the construct validity sample (Fig. 1).

Statistical Methods

To evaluate whether patients in the test-retest reliability sample who completed the retest systematically differed from patients who did not complete the retest, we performed a nonresponder analysis. In this analysis, we classified nonresponders as patients who did not complete the retest in the predetermined time, and responders were classified as patients who completed the retest. The sociodemographics of responders and nonresponders were compared using independent sample t-tests for continuous data and chi-square tests for dichotomous or categorical data. A p value < 0.05 was considered statistically significant. The sociodemographics of responders and nonresponders were highly similar; the only difference was whether the dominant side was treated, this was the case in 53% of the responders and in 33% of the nonresponders (p = 0.03) (Supplementary Table 1; supplemental materials are available with the online version of *CORR*®).

Test-retest reliability was evaluated using the weighted kappa and ICCs for question 1 of the Satisfaction with Treatment Result Questionnaire, and Cohen's kappa was used for question 2. We also evaluated test-retest reliability using Cohen's kappa for dichotomized modifications of question 1 because these might be used in logistic regression models in future research. For this, the 5-point Likert scale will be split into "satisfied" and "dissatisfied" using two classifications, with the answering options of "poor," "moderate," and "fair" attributed to "dissatisfied" in the first classification and only "poor" and "moderate" attributed to "dissatisfied" in the second classification. For the weighted kappa determination, we used quadratic weights, implying that misclassification between adjacent categories is less problematic than those between more distant categories. The greater the distance, the larger the penalty for misclassification [23,24]. For instance, a deviation from "good" to "poor" gets more weight than a deviation from "good" to "fair." Weighted kappa and Cohen's kappa scores can range from -1 to 1, where <0 indicates no agreement, 0.01 to 0.20 is none to slight, 0.21 to 0.40 is fair, 0.41 to 0.60 is moderate, 0.61 to 0.80 is substantial, and 0.81 to 1.00 is almost-perfect agreement [23].

ICC values were calculated using a two-way mixed-effects model [25]. ICC values range from 0 to 1, where 1 is perfect reliability, 0.90 to 0.99 is very high reliability; 0.70 to 0.89 indicates high reliability; 0.50 to 0.69 represents moderate reliability; 0.26 to 0.49 is low reliability, and 0.00 to 0.25 indicates little, if any, reliability [26-28]. We also calculated the percentage of absolute agreement between the primary test and the retest for both questions and both dichotomized variants to examine the absolute proportion of overlap between the primary test and the re-test. The absolute percentage agreement was considered high if it exceeded 75%, moderate if it was between 40% and 75%, and low if it was less than 40%.

Construct validity was evaluated using hypotheses testing, following the guidelines of the Consensus-based Standards for the Selection of Health Measurement Instruments [29]. Construct validity was defined as "the degree to which the scores of a measurement instrument are consistent with the hypotheses, with regard to internal relationships, relationships with scores of other instruments, or differences between relevant groups"[29].

We formed eight hypotheses prior to the analysis, with a specific and clearly defined direction, magnitude, and rationale (Table 4). First, we hypothesized that there was a strong association between Question 1 and Question 2 of the Satisfaction with Treatment Result Questionnaire, derived from the rationale that satisfaction with result may dictate the decision to undergo the treatment again in the future. Also, we hypothesized that the level of satisfaction would have at least a moderate correlation with pain and function levels, as, logically, these may determine one's level of satisfaction. Furthermore, we hypothesized that one's level of recommendation would moderately correlate with level of satisfaction, as the degree of recommendation may, among other things, be influenced by satisfaction with the treatment result. Lastly, we hypothesized that there would be a strong correlation with satisfaction with the hand, as this construct may overlap with satisfaction with treatment result. For each possible outcome, we also defined the interpretation before the analysis. All authors agreed with the eight independent hypotheses before analysis. We considered each hypothesis with equal weight. To test the hypotheses, we calculated the Spearman rho correlation coefficients between question 1 of the Satisfaction with Treatment Result Questionnaire and question 2, VAS pain during physical load, VAS function, VAS satisfaction with the hand, the NPS, MHQ pain, MHQ general hand function, and MHQ satisfaction. The Spearman correlation coefficients were interpreted as follows: 0.00 to 0.19 is a very weak correlation, 0.20.to 0.39 is a weak correlation, 0.40 to 0.69 is a moderate correlation, 0.70 to 0.89 is a strong correlation, and 0.90 to 1 is a very strong correlation [30,31]. Confirmation of ≥ 80% of the hypotheses was considered good-to-excellent construct validity [29].

All analyses were performed using R Statistical Programming, version 3.3.4 (R Project for Statistical Computing).

Results

Test-retest Reliability

We found high reliability and almost-perfect agreement for test-retest reliability using the 5-point Likert scale of the Satisfaction with Treatment Result Questionnaire, with an ICC value of 0.86 (95% CI 0.81 to 0.89) and a weighted kappa of 0.86 (95% CI 0.80 to 0.91), respectively (Table 3). The distribution of answers at the primary test and the retest were highly similar (Fig. 2A) and most deviations were one step up or down compared to the primary test (Fig. 2B). The first dichotomized variant of question 1, with "poor," "moderate," and "fair" attributed to "dissatisfied," demonstrated an absolute percentage of agreement of 87% and a kappa score of 0.73 (95% CI 0.62 to 0.83), indicating substantial agreement. The second dichotomized variant of question 1, with "poor" and "moderate" attributed to "dissatisfied," demonstrated an absolute percentage of agreement of 81% and a kappa score of 0.57 (95% CI 0.45 to 0.69), indicating moderate agreement. When patients were asked about their willingness to undergo treatment again, we found a kappa score of 0.81 (95% CI 0.70 to 0.92), indicating almost-perfect agreement (94%).

We included both Intraclass Correlation (ICC) and Weighted kappa as the ICC indicates the reliability and the Weighted kappa indicates the weighted agreement of the retest relative to the primary test.

Table 3. Overview of the outcomes of test-retest reliability

Question of Satisfaction with Treatment Result Questionnaire	ICC (95% CI)	Weighted kappa (95% CI)	Kappa (95% CI)	Absolute agreement
Question 1: level of satisfaction measured with a five-point Likert scale	0.86 (0.81-0.89)	0.86 (0.80-0.91)		70%
Question 1: level of satisfaction measured with dichotomized variant 1 (“poor,” “moderate,” and “fair” attributed to “dissatisfied”)			0.73 (0.62-0.83	87%
Question 1: level of satisfaction measured with dichotomized variant 2 (“poor” and “moderate” attributed to “dissatisfied”)			0.57 (0.45-0.69)	81%
Question 2: willingness to undergo the treatment again			0.81 (0.70-0.92)	94%

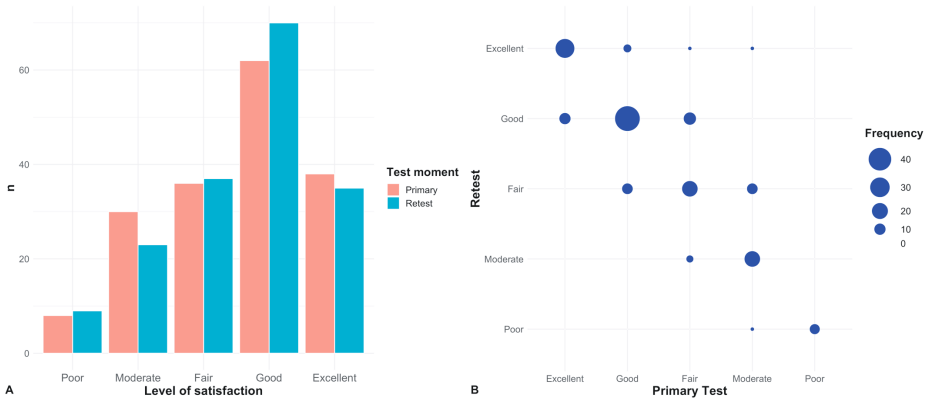


Fig. 2 (A) This bar plot indicates the distribution of question 1 of the Satisfaction with Treatment Result Questionnaire at the primary test moment and the retest. (B) This balloon plot indicates the degree of deviation between the primary test and the retest of question 1 of the Satisfaction with Treatment Result Questionnaire. In this plot, the primary test is displayed on the x axis, while the retest score is displayed on the y axis. The size of the dots indicates the number of patients.

Table 4. Hypotheses for testing the construct validity of satisfaction with the treatment result and associated conclusions

Hypothesis for Question 1 of the Satisfaction with Treatment Result Questionnaire	Rationale	Hypothesized Spearman's rho[a]	Conclusions if correlation coefficient is lower than hypothesized	Conclusions if correlation coefficient is higher than hypothesized	Actual correlation	Hypothesis confirmed?
Entire construct validity sample (n=3750)						
1. There will be at least a strong association between Question 1 and Question 2 of the Satisfaction with Treatment Result Questionnaire	Satisfaction with result may dictate the decision to undergo the treatment again in the future	≥ 0.7	Very weak, weak, or moderate correlation. Constructs may differ.	Strong or very strong correlation. Constructs may be the same or overlap.	0.44*	No
2. There will be at least a moderate association with VAS pain during physical load	Residual pain may correlate with satisfaction with the treatment result	≥ 0.4	Very weak or weak correlation. Constructs are very different.	Moderate or strong correlation with a different construct.	0.67*	Yes
3. There will be at least a moderate association with VAS function	Level of function may correlate with satisfaction with treatment result	≥ 0.4	Very weak or weak correlation. Constructs are very different.	Moderate or strong correlation with different construct.	0.61*	Yes
4. There will be at least a strong association with VAS satisfaction with the hand	Satisfaction with treatment result and satisfaction with the hand might measure an overlapping construct	≥ 0.7	Very weak, weak, or moderate correlation, although theoretically, there is overlap of constructs.	Strong or very strong correlation with, theoretically, an overlapping construct.	0.73*	Yes

Table 4. Hypotheses for testing the construct validity of satisfaction with the treatment result and associated conclusions *(continued)*

Hypothesis for Question 1 of the Satisfaction with Treatment Result Questionnaire	Rationale	Hypothesized Spearman's rho[a]	Conclusions if correlation coefficient is lower than hypothesized	Conclusions if correlation coefficient is higher than hypothesized	Actual correlation	Hypothesis confirmed?
5. There will be at least a moderate association with the NPS	Satisfaction with the treatment result may correlate to the NPS because the degree of recommendation may, among other things, be influenced by satisfaction with the treatment result.	≥ 0.4	Very weak or weak correlation. Constructs are very different.	Moderate or strong correlation with different construct.	0.43*	Yes
Subset of construct validity sample with MHQ data (n=2306)						
6. There will be at least a moderate association with MHQ pain subscale	Residual pain may correlate with satisfaction with the treatment result	≥ 0.4	Very weak or weak correlation. Constructs are very different.	Moderate or strong correlation with a different construct.	0.57*	Yes
7. There will be at least a moderate association with MHQ overall hand function subscale	Level of function may correlate with satisfaction with treatment result	≥ 0.4	Very weak or weak correlation. Constructs are very different.	Moderate or strong correlation with different construct.	0.51*	Yes
8. There will be at least a strong association with MHQ satisfaction with the hand	Satisfaction with treatment result and satisfaction with the hand might measure an overlapping construct	≥ 0.7	Very weak, weak, or moderate correlation, although theoretically, there is overlap of constructs.	Strong or very strong correlation with, theoretically, an overlapping construct.	0.71*	Yes

[a]Spearman's correlation coefficients are interpreted as follows: 0.00-0.19 = very weak correlation; 0.20-0.39 = weak correlation; 0.40-0.69 = moderate correlation; 0.70-0.89 = strong correlation; and 0.90-1 = very strong correlation. NPS = Net Promotor Score, VAS = Visual Analogue Scale, MHQ = Michigan Hand outcomes Questionnaire * $p<0.001$.

Construct Validity

The Satisfaction with Treatment Result Questionnaire demonstrated good-to-excellent construct validity in this study. Of the eight hypotheses we tested, seven confirmed construct validity (Table 4). In the confirmed hypotheses, there was a moderate-to-strong correlation with VAS pain, VAS function, NPS, MHQ pain, and MHQ general hand function (Spearman rho ranging from 0.43 to 0.67; all $p < 0.001$) and a strong to very strong correlation with VAS satisfaction and MHQ satisfaction (Spearman rho of 0.73 and 0.71; both $p < 0.001$). Only hypothesis 1 was rejected, as we found only a moderate correlation between question 1 and question 2 of the Satisfaction with Treatment Result Questionnaire (Spearman rho of 0.44; $p < 0.001$).

Discussion

Satisfaction with treatment result is widely used and is considered an essential and patient-centered outcome domain [11,18]. Before this study, there were doubts on reliability and validity of measures of satisfaction with the treatment result [12].

In this study, we found that the Satisfaction with Treatment Result Questionnaire had good-to-excellent construct validity and very high test-retest validity in two large samples of patients with hand and wrist conditions. Our findings indicate that the Satisfaction with Treatment Result Questionnaire is a reliable and valid instrument that can safely be used in daily practice and clinical research for evaluating patients' satisfaction with their treatment result after treatment for a hand or wrist condition.

Limitations

A limitation of the observational design of this study is that a substantial proportion of patients did not respond, although our nonresponder analysis indicated that there were very few differences between responders and nonresponders. Hence, we are confident that this did not influence our results. An additional limitation is that we evaluated construct validity in the absence of a gold standard. Future research should investigate how to address this. Additionally, although not in the scope of this study, another limitation is that we did not study other important psychometric properties of the Satisfaction with Treatment Result Questionnaire, including responsiveness and other aspects of validity such as content validity [12]. Also, the psychometric properties of this measure in other study populations are still unknown.

Test-retest Reliability

Our study shows that satisfaction with a treatment result can reliably be measured using a one-question, 5-point Likert scale. Because we did not find any other studies investigating the psychometrical properties of the Satisfaction with Treatment Result Questionnaire in hand and wrist conditions, we cannot compare our findings with previous studies. However, Ring and Leopold. [12] questioned the validity and reliability of assessing satisfaction with treatment results using a PROM, owing to within-person variation in

pretreatment expectations and environmental and psychological factors. Although variation in these constructs exists among patients with hand and wrist conditions [32-34], our study shows that satisfaction with a treatment result can be measured reliably using a standardized PROM such as ours. This is supported by our finding that, if deviations between test-retest measurements occurred, these deviations were, in almost all instances, only in one level on the 5-point Likert scale.

We found that a dichotomized variant of a patient's level of satisfaction, with "poor," "moderate," and "fair" attributed to "dissatisfied," yielded substantial agreement, while the other variant yielded only moderate agreement. Although the agreement decreased when dichotomizing outcomes, which is often suboptimal due to loss of data, use of the first variant may be useful, for example, when aiming to use logistic regression models to explain the variance in levels of satisfaction with a treatment result.

Construct Validity

We found good-to-excellent construct validity of the Satisfaction with Treatment Result Questionnaire in this study as seven of the eight hypotheses we tested were confirmed. However, it should be noted that a gold standard for measuring satisfaction with treatment result is absent. Additionally, although the VAS satisfaction and MHQ satisfaction evaluate satisfaction with one's hand and not satisfaction with treatment result, there may be circular reasoning. Future studies of construct validity may incorporate additional measures, such as the Global Rating of Change Score.

A remarkable finding in this study is that a patient's willingness to undergo the treatment again under similar circumstances (question 2 of the Satisfaction with Treatment Result Questionnaire) was only moderately associated with his or her level of satisfaction with the treatment result (question 1). An explanation for this finding may be that a patient might not be completely satisfied with the treatment result but has improved enough to consider the treatment again under similar circumstances (or vice versa). This suggests that these two questions measure different constructs, and future research should investigate how these two constructs relate. Furthermore, the influence of one's psychological mindset (including aspects such as anxiety or depression) and other factors on levels of satisfaction and willingness to undergo a treatment again should be further explored [12]. Also, future research may investigate which components form the construct of satisfaction with treatment result from both a patient and clinician perspective to optimize validity in measures of satisfaction with treatment result.

Conclusion

In this study, the Satisfaction with Treatment Result Questionnaire had good-to-excellent construct validity and very high test-retest validity in two large samples of patients with hand and wrist conditions. Satisfaction with treatment result can be measured safely in daily practice and clinical research using these questions in striving for patient-centered

care and value-based health care. Future research should investigate other psychometric properties such as responsiveness or content validity, other tools such as the International Consortium for Health Outcomes Measurement satisfaction with treatment result questionnaire, as well as independent predictors of variation in satisfaction with the treatment result.

Acknowledgments

We thank all patients who completed questionnaires as part of their clinical care and who agreed that their data could be anonymously used for the present study. In addition, we thank the members of the Hand-Wrist Study Group, caregivers, and personnel of Xpert Clinic, Handtherapie Nederland, and Equipe Zorgbedrijven for assisting in the routine outcome measurements that are the basis for this study.

Group Authors

Members of the Hand-Wrist Study Group include R. A. M. Blomme, B. J. R. Sluijter, D. J. J. C. van der Avoort, A. Kroeze, J. Smit, J. Debeij, E. T. Walbeehm, G. M. van Couwelaar, G. M. Vermeulen, J. P. de Schipper, J. F. M. Temming, J. H. van Uchelen, H. L. de Boer, K. P. de Haas, K. Harmsen, O. T. Zöphel, R. Feitz, J. S. Souer, R. Koch, S. E. R. Hovius, T. M. Moojen, X. Smit, R. van Huis, P. Y. Pennehouat, K. Schoneveld, Y. E. van Kooij, J. J. Veltkamp, A. Fink, W. A. de Ridder, H. P. Slijper, R. W. Selles, J. T. Porsius, R. M. Wouters, J. Tsehaie, R. Poelstra, M. C. Jansen, M. J. W. van der Oest, L. Hoogendam, J. S. Teunissen, J. Dekker, M. Jansen-Landheer, M. ter Stege, J. M. Zuidam, C. A. van Nieuwenhoven, L. S. Duraku, C. Hundepool, B. van der Heijden, and J. W. Colaris.

References

1. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477-2481.
2. Barry MJ, Edgman-Levitan S. Shared Decision Making — The Pinnacle of Patient-Centered Care. *New Engl J Med.* 2012;366(9):780-781.
3. Basch E. Patient-Reported Outcomes — Harnessing Patients' Voices to Improve Clinical Care. *New Engl J Med.* 2017;376(2):105-108.
4. Mandl LA, Galvin DH, Bosch JP, George CC, Simmons BP, Axt TS, Fossel AH, Katz JN. Metacarpophalangeal arthroplasty in rheumatoid arthritis: What determines satisfaction with surgery? *J Rheumatol.* 2002;29(12):2488-2491.
5. van Oosterom FJ, Ettema AM, Mulder PG, Hovius SE. Impairment and disability after severe hand injuries with multiple phalangeal fractures. *J Hand Surg Am.* 2007;32(1):91-95.
6. MacDermid JC, Roth JH, Raj Rampersaud Y, Bain GI. Trapezial arthroplasty with silicone rubber implantation for advanced osteoarthritis of the trapeziometacarpal joint of the thumb. *Can J Surg.* 2003;46(2):103-110.
7. Karnezis IA, Fragkiadakis EG. Association between objective clinical variables and patient-rated disability of the wrist. *J Bone Joint Surg Br.* 2002;84(7):967-970.
8. Goldhahn J, Angst F, Simmen BR. What counts: outcome assessment after distal radius fractures in aged patients. *J Orthop Trauma.* 2008;22(8 Suppl):S126-130.
9. Bleich SN, Ozaltin E, Murray CK. How does satisfaction with the health-care system relate to patient experience? *Bull World Health Organ.* 2009;87(4):271-278.
10. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev.* 2013;70(4):351-379.
11. Marks M, Herren DB, Vliet Vlieland TPM, Simmen BR, Angst F, Goldhahn J. Determinants of patient satisfaction after orthopedic interventions to the hand: A review of the literature. *J Hand Ther.* 2011;24(4):303-312.
12. Ring D, Leopold SS. Editorial-Measuring Satisfaction: Can It Be Done? *Clin Orthop Relat Res.* 2015;473(10):3071-3073.
13. Vranceanu AM, Ring D. Factors associated with patient satisfaction. *J Hand Surg Am.* 2011;36(9):1504-1508.
14. Graham B, Green A, James M, Katz J, Swiontkowski M. Measuring patient satisfaction in orthopaedic surgery. *The Journal of bone and joint surgery American volume.* 2015;97(1):80-84.
15. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis care & research.* 2011;63 Suppl 11:S240-252.
16. MacDermid JC, Wessel J, Humphrey R, Ross D, Roth JH. Validity of self-report measures of pain and disability for persons who have undergone arthroplasty for osteoarthritis of the carpometacarpal joint of the hand. *Osteoarthritis Cartilage.* 2007;15(5):524-530.

17. Poole J. Measures of hand function: Arthritis Hand Function Test (AHFT), Australian Canadian Osteoarthritis Hand Index (AUSCAN), Cochin Hand Function Scale, Functional Index for Hand Osteoarthritis (FIHOA), Grip Ability Test (GAT), Jebsen Hand Function Test (JHFT), and Michigan Hand Outcomes Questionnaire (MHQ). *Arthritis care & research.* 2011;63 Suppl 11:S189-S199.

18. Wouters RM, Jobi-Odeneye AO, De la Torre A, Joseph A, ICHOM Hand and Wrist Working Group, Hovius SER. Hand and Wrist Conditions Data Collection Reference Guide Version 1.0.0. Published 2020. Accessed April 1st, 2021.

19. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453-1457.

20. Selles RW, Wouters RM, Poelstra R, van der Oest MJW, Porsius JT, Hovius SER, Moojen TM, van Kooij Y, Pennehouat PY, van Huis R, Vermeulen GM, Feitz R, Slijper HP. Routine Health Outcome Measurement: Development, Design, and Implementation of the Hand and Wrist Cohort. *Plast Reconstr Surg.* 2020;146(2):343-354.

21. Breckenridge K, Bekker HL, Gibbons E, van der Veer SN, Abbott D, Briançon S, Cullen R, Garneata L, Jager KJ, Lønning K, Metcalfe W, Morton RL, Murtagh FE, Prutz K, Robertson S, Rychlik I, Schon S, Sharp L, Speyer E, Tentori F, Caskey FJ. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. *Nephrol Dial Transplant.* 2015;30(10):1605-1614.

22. Manary MP, Boulding W, Staelin R, Glickman SW. The patient experience and health outcomes. *N Engl J Med.* 2013;368(3):201-203.

23. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70(4):213-220.

24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.

25. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;15(2):155-163.

26. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420-428.

27. Mitani AA, Freer PE, Nelson KP. Summary measures of agreement and association between many raters' ordinal classifications. *Ann Epidemiol.* 2017;27(10):677-685.e674.

28. Fleiss JL. Reliability of Measurement. In: *The Design and Analysis of Clinical Experiments.*1999:1-32.

29. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745.

30. Streiner D, Norman G, Cairney J. Health measurement scales: a practical guide to their development and use (5th edition). *Aust N Z J Public Health.* 2016;40(3):294-295.

31. Balkau B. *Practical Statistics for Nursing and Health Care. In: European Diabetes Nursing.* Vol 3: Taylor & Francis; 2006: 27-27.

32. Tsehaie J, Spekreijse KR, Wouters RM, Feitz R, Hovius SER, Slijper HP, Selles RW. Predicting Outcome After Hand Orthosis and Hand Therapy for Thumb Carpometacarpal Osteoarthritis: A Prospective Study. *Arch Phys Med Rehabil.* 2019;100(5):844-850.

33. Wouters RM, Vranceanu AM, Slijper HP, Vermeulen GM, van der Oest MJW, Selles RW, Porsius JT. Patients With Thumb-base Osteoarthritis Scheduled for Surgery Have More Symptoms, Worse Psychological Profile, and Higher Expectations Than Nonsurgical Counterparts: A Large Cohort Analysis. *Clin Orthop Relat Res.* 2019;477(12):2735-2746.

34. Wouters RM, Slijper HP, Esteban Lopez L, Hovius SER, Selles RW. Beneficial Effects of Nonsurgical Treatment for Symptomatic Thumb Carpometacarpal Instability in Clinical Practice: A Cohort Study. *Arch Phys Med Rehabil.* 2020;101(3):434-441.

Supplementary Table 1: Non-responder analysis of the first sample.

Variables	Responders	Non-responders	P-value
Participants, n	174	113	
Age (mean (SD))	56.4 (12.8)	53.5 (16%)	0.089
Sex = male, n (%)	72 (41%)	40 (35%)	0.373
Dominant side treated = yes, n (%)	93 (53%)	37 (33%)	**0.028**
Second opinion = no, n (%)	167 (96%)	110 (97%)	0.773
Type of work, n (%)			0.573
Unemployed	61 (35%)	45 (40%)	
Light physical labor	51 (29%)	28 (25%)	
Moderate physical labor	53 (31%)	31 (27%)	
Heavy physical labor	9 (5%)	9 (8%)	
Symptom duration in months, mean (SD)	22.6 (45%)	17.2 (25%)	0.241
Measurement track, n (%)			0.577
Thumb regular	12 (7%)	15 (13%)	
Thumb extended	15 (9%)	9 (8%)	
Wrist regular	31 (18%)	23 (20%)	
Wrist extended	18 (10%)	8 (7%)	
Finger regular	33 (19%)	17 (15%)	
Finger extended	8 (5%)	3 (3%)	
Dupuytren's	25 (14%)	19 (17%)	
Nerve (de)compression	32 (18%)	19 (17%)	
Primary test of satisfaction with treatment result: question 1, n (%)			0.855
Excellent	38 (22%)	27 (24%)	
Good	62 (36%)	42 (37%)	
Fair	36 (21%)	18 (16%)	
Moderate	30 (17%)	19 (17%)	
Poor	8 (5%)	7 (6%)	
Primary test of satisfaction with treatment result: question 2 = no, n (%)	34 (20%)	20 (18%)	0.814

3

WHICH FACTORS ARE ASSOCIATED WITH SATISFACTION WITH TREATMENT RESULTS IN PATIENTS WITH HAND AND WRIST CONDITIONS? A LARGE COHORT ANALYSIS

de Ridder WA, Wouters RM, Hoogendam L, Vermeulen GM, Slijper HP, Selles RW, the Hand-Wrist Study Group; 2022; Clinical Orthopaedics and Related Research.

Abstract

Background

Satisfaction with treatment results is an important outcome domain in striving for patient-centered and value-based healthcare. Although numerous studies have investigated factors associated with satisfaction with treatment results, most studies used relatively small samples. Additionally, many studies have only investigated univariable associations instead of multivariable associations; to our awareness, none have investigated the independent association of baseline sociodemographics, quality of life, improvement in pain and function, experiences with healthcare delivery, and baseline measures of mental health with satisfaction with treatment results.

Questions/purposes

(1) What factors are independently associated with satisfaction with treatment results at 3 months post-treatment in patients treated for common hand and wrist conditions? (2) What factors are independently associated with the willingness to undergo the treatment again at 3 months post-treatment in patients treated for common hand and wrist conditions? Among the factors under study were baseline sociodemographics, quality of life, improvement in pain and function, experiences with healthcare delivery, and baseline measures of mental health.

Methods

Between August 2018 and May 2020, we included patients who underwent carpal tunnel release, nonsurgical or surgical treatment for thumb-base osteoarthritis, trigger finger release, limited fasciectomy for Dupuytren's contracture, or nonsurgical treatment for midcarpal laxity in one of the 28 centers of Xpert Clinics in the Netherlands. We screened 5859 patients with complete sociodemographics and data at baseline. Thirty-eight percent (2248 of 5859) of these patients had complete data at 3 months. Finally, participants were eligible for inclusion if they provided a relevant answer to the three patient-reported experience measures (PREM) items. A total of 424 patients did not do this because they answered "I don't know" or "not applicable" to a PREM item, leaving 31% (1824 of 5859) for inclusion in the final sample. A validated Satisfaction with Treatment Result Questionnaire was administered at 3 months, which identified the patients' level of satisfaction with treatment results so far on a 5-point Likert scale (research question 1, with answers of good, excellent, poor, moderate, or fair) and the patients' willingness to undergo the treatment again under similar circumstances (research question 2, with answers of yes or no). A hierarchical logistic regression model was used to identify whether baseline sociodemographic, change in outcome (patient-reported outcome measures for quality of life, hand function, and pain), baseline measures of mental health (including treatment credibility [the extent to which a patient attributes credibility to a treatment] and expectations, illness perception, pain catastrophizing, anxiety and depression), and patient-reported experience measures (PREMs) were associated with

each question of the Satisfaction with Treatment Result Questionnaire at 3 months post-treatment. We dichotomized our first question into good and excellent, which were considered more satisfied, and poor, moderate, and fair were considered less satisfied. After dichotomization, 57% (1042 of 1824) of patients were classified as more satisfied with the treatment results.

Results

The following variables were independently associated with satisfaction with treatment results, with an area under the curve of 0.82 (95% confidence interval 0.80 to 0.84) (arranged from the largest to the smallest standardized odds ratio): greater decrease in pain during physical load (SOR 2.52 [95% CI 2.18 to 2.92]; $p < 0.001$), patient's positive experience with the explanation of the pros and cons of the treatment (determined with the question: "Have you been explained the pros and cons of the treatment or surgery?") (SOR 1.83 [95% CI 1.41 to 2.38]; $p < 0.001$), greater improvement in hand function (SOR 1.76 [95% CI 1.54 to 2.01]; $p < 0.001$), patients' positive experience with the advice for at home (determined with the question: "Were you advised by the healthcare providers on how to deal with your illness or complaints in your home situation?") (SOR 1.57 [95% CI 1.21 to 2.04]; $p < 0.001$), patient's better personal control (determined with the question: "How much control do you feel you have over your illness?") (SOR 1.24 [95% CI 1.1 to 1.40]; $p < 0.001$), patient's more positive treatment expectations (SOR 1.23 [95% CI 1.04 to 1.46]; $p = 0.02$), longer expected illness duration by the patient (SOR 1.2 [95% CI 1.04 to 1.37]; $p = 0.01$), a smaller number of symptoms the patient saw as part of the illness (SOR 0.84 [95% CI 0.72 to 0.97]; $p = 0.02$), and less concern about the illness the patient experiences (SOR 0.84 [95% CI 0.72 to 0.99]; $p = 0.04$). For willingness to undergo the treatment again, the following variables were independently associated with an AUC of 0.81 (95% CI 0.78 to 0.83) (arranged from the largest to the smallest standardized OR): patient's positive experience with the information about the pros and cons (determined with the question: "Have you been explained the pros and cons of the treatment or surgery?") (SOR 2.05 [95% CI 1.50 to 2.8]; $p < 0.001$), greater improvement in hand function (SOR 1.80 [95% CI 1.54 to 2.11]; $p < 0.001$), greater decrease in pain during physical load (SOR 1.74 [95% CI 1.48 to 2.07]; $p < 0.001$), patient's positive experience with the advice for at home (determined with the question: "Were you advised by the healthcare providers on how to deal with your illness or complaints in your home situation?") (SOR 1.52 [95% CI 1.11 to 2.07]; $p = 0.01$), patient's positive experience with shared decision-making (determined with the question: "Did you decide together with the care providers which care or treatment you will receive?") (SOR 1.45 [95% CI 1.06 to 1.99]; $p = 0.02$), higher credibility the patient attributes to the treatment (SOR 1.44 [95% CI 1.20 to 1.73]; $p < 0.001$), longer symptom duration (SOR 1.27 [95% CI 1.09 to 1.52]; $p < 0.01$), and patient's better understanding of the condition (SOR 1.17 [95% CI 1.01 to 1.34]; $p = 0.03$).

Conclusion

Our findings suggest that to directly improve satisfaction with treatment results, clinicians might seek to: (1) improve the patient's experience with healthcare delivery, (2) try to influence illness perception, and (3) boost treatment expectations and credibility. Future research should confirm if these suggestions are valid and perhaps also investigate whether satisfaction with treatment results can be predicted (instead of explained, as was done in this study). Such prediction models, as well as other decision support tools that investigate patient-specific needs, may influence experience with healthcare delivery, expectations, or illness perceptions, which in turn may improve satisfaction with treatment results.

Introduction

Satisfaction with treatment results is an important outcome domain in striving for patient-centered and value-based healthcare. In these frameworks, the patient is central, and the aim is to achieve high value at low cost [1-6]. After all, is there value in a technically perfect surgical procedure, with no complications and excellent objective outcomes afterwards, if the patient is not satisfied with the treatment results? Although recognized as an important outcome domain [7], the interpretation of satisfaction with treatment results is difficult, and there are doubts about the face validity of questionnaires to measure satisfaction with treatment results [8]. However, the Satisfaction with Treatment Result Questionnaire has a good-to-excellent construct validity and a very high test-retest reliability [9], and we believe it is reasonable to use it in a study exploring this topic.

Several studies have investigated factors associated with satisfaction with treatment results [10-30]. Marks et al. [10] found associations between satisfaction and pain and symptoms, activities of daily living or function, aesthetics, embodiment, strength, ROM, fulfillment of expectations, deformity, workers compensation, and length of follow-up. Additionally, strong associations have been found between satisfaction and better patient-reported experience measures (PREMs), such as the provision of general and treatment information, and with physician communication and shared decision-making [11-16]. Furthermore, the relationship with the surgeon, particularly perceived empathy, is a driver of satisfaction with treatment results [17-20]. Associations with several measures of mental health have also been found. For instance, higher preoperative pain catastrophizing is associated with lower satisfaction after hand surgery [21,22], and more symptoms of depression are associated with greater dissatisfaction after carpal tunnel release [23]. There is no consensus on the association between treatment expectations and satisfaction with treatment results; several authors suggested that higher expectations may lead to lower satisfaction [24-26], whereas other studies found a reverse association [13,27-30].

Although the aforementioned studies investigated factors associated with patient satisfaction with treatment results, most studies used relatively small samples or used a univariable approach instead of a multivariable approach. Therefore, the independent association of baseline sociodemographics, quality of life, improvement in pain and function, experiences with healthcare delivery, and baseline measures of mental health with satisfaction with treatment results is still unclear. More knowledge on independent factors that are associated with satisfaction with treatment results may help clinicians to directly improve satisfaction with treatment results, as well as inform future studies aiming to develop interventions that improve satisfaction with treatment results.

Therefore, we asked: (1) What factors are independently associated with satisfaction with treatment results at 3 months post-treatment in patients treated for common hand and wrist conditions? (2) What factors are independently associated with the willingness to

undergo the treatment again at 3 months post-treatment in patients treated for common hand and wrist conditions? Among the studied factors were baseline sociodemographics, quality of life, improvement in pain and function, experiences with healthcare delivery, and baseline measures of mental health.

Patients and Methods

Study Design

This was a cohort study using a longitudinally maintained, population-based sample of patients with hand and wrist conditions from the Hand Wrist Study Group cohort, reported following the Strengthening the Reporting of Observational Studies in Epidemiology guidelines [31].

Setting

Data collection using GemsTracker electronic data capture tools (GemsTracker 2020) was part of usual care and occurred between August 2018 and May 2020 at Xpert Clinics. The start date of the current PREM determined the start date of the study. All data were digitally collected using GemsTracker, a secure internet-based application for distributing questionnaires and forms during clinical research and quality registrations. Xpert Clinics comprises 28 clinics for hand surgery and therapy in The Netherlands. Twenty-three surgeons who have been certified by the Federation of European Societies for Surgery of the Hand and more than 150 hand therapists are employed at our treatment centers. At Xpert Clinics, treatment outcomes are evaluated in measurement tracks, each of which consists of treatments with similar relevant outcome domains and timepoints. After a diagnosis is registered during the first consultation, a measurement track is automatically activated, and patient-reported outcome measure forms are emailed to the patient. Details of this procedure have been published [32].

Participants

Participants were eligible for inclusion if they were adults who completed all relevant questionnaires. We included patients who underwent one of the following common treatments: trigger finger release (23% [423 of 1824]), limited fasciectomy (17% [307 of 1824]), trapeziectomy with or without ligament reconstruction tendon interposition for thumb base osteoarthritis (12% [213 of 1824]), carpal tunnel release (29% [521 of 1824]), hand therapy for midcarpal laxity (2% [35 of 1824]), and hand therapy for thumb base osteoarthritis (18% [325 of 1824]) (Table 1). Because the aim of this study was to investigate which factors explain satisfaction with treatment results in a general population of patients treated for hand and wrist disorders, we selected the largest pathology of each of the six largest measurement tracks from our cohort [32]. Patients who underwent operative treatment were assessed at 3 months postoperatively, and patients who underwent nonoperative treatment were assessed 3 months after treatment was initiated.

Table 1. Characteristics at baseline of all included patients (n = 1824)

Variable	Value
Age in years	59 ± 11
Sex (Male)	39 (704)
Second opinion	2 (42)
Recurrence (Yes)	8 (146)
Hand dominance	
Right	88 (1607)
Left	8 (153)
Both	4 (64)
Dominant hand treated	49 (902)
Symptom duration in months median (interquartile range)	12 (6-24)
Workload	
Not employed	40 (734)
Light	27 (492)
Moderate	23 (427)
Severe	9 (171)
BMI in kg/m^2	27 ± 5
Smoking (No)	86 (1571)

Data presented as mean ± SD or % (n), unless otherwise noted.

We screened 5859 patients with complete sociodemographics and data at baseline. Thirty-eight percent (2248 of 5859) of patients had complete data at 3 months. Finally, participants were eligible for inclusion if they provided a relevant answer to the three PREM items. A total of 424 patients did not do this because they answered "I don't know" or "not applicable" to a PREM item, leaving 31% (1824 of 5859) for inclusion in the final sample (Fig. 1). There were no additional exclusion criteria.

To assess potential selection bias, we compared responder and nonresponder demographics and measures of mental health, using the standardized mean difference as an indication of imbalance [33]. Nonresponders were defined as patients who did not complete questionnaires at 3 months or did not provide a relevant answer to a PREM item. Responders were defined as patients who completed all relevant questionnaires at baseline and at 3 months. Responders and nonresponders both received treatment and remained in care. Besides difference in treatment type (Standardized Mean Difference 0.26), all standardized mean difference values were < 0.2, indicating that the magnitude of the standardized mean difference was even smaller than that defined as small by Cohen [34] (Supplementary Table 1; supplemental materials are available with the online version of *CORR*®). Additionally, we conducted a Little test (p = 0.27), which supported the idea that nonresponders could be considered missing at random [35-37].

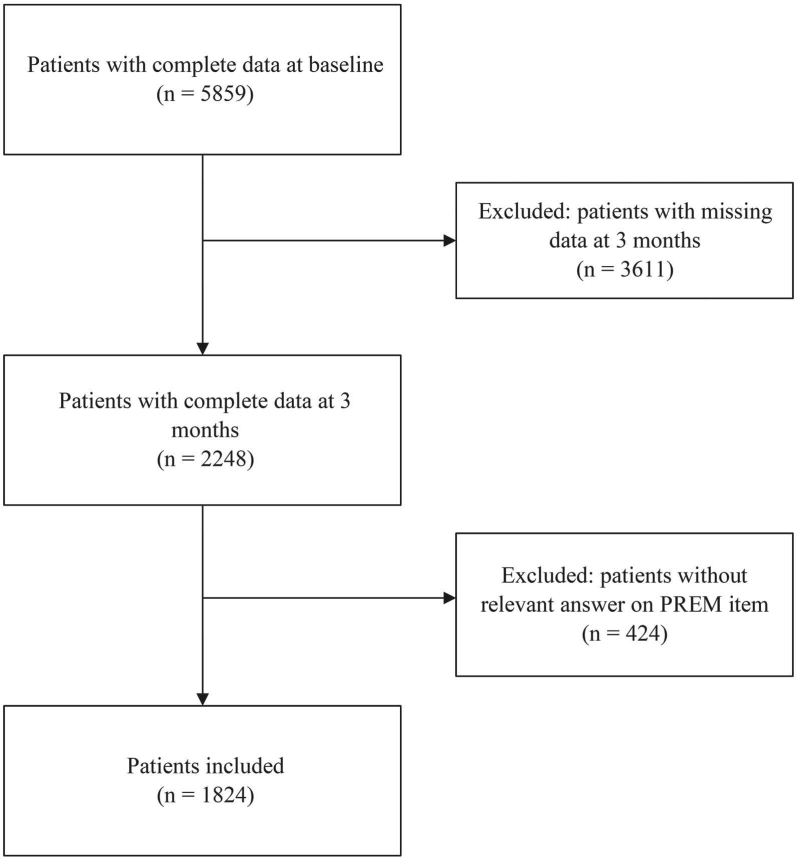


Fig. 1 This flowchart illustrates the patient selection for this study.

Variables and Measurements

The primary outcomes in this study were the two questions of the Satisfaction with Treatment Result Questionnaire at 3 months after the start of treatment. The first question evaluates patients' satisfaction with treatment results on a 5-point Likert scale (answering options: poor, moderate, fair, good, and excellent). In the second question, patients indicated whether they would undergo the same procedure again under similar circumstances (yes or no). The Satisfaction with Treatment Result Questionnaire has a good-to-excellent construct validity and very high test-retest validity [9].

We classified variables we investigated as potentially associated with satisfaction into four categories: sociodemographic, clinical patient-reported outcome measures, measures of mental health, and PREMs.

Sociodemographic characteristics included sex (we report sex, not gender, as our data comes from the Dutch Citizen Service Administration, so we did not want to make any

unsupported assumptions on gender), age, occupational status (unemployed or light, medium, or heavy physical labor), whether the patient visited the clinic for a second opinion, self-reported duration of symptoms (in months), whether the dominant hand was treated, and whether the disease was recurrent (measured by the question: “Have been treated for the same disease before?”; the answer yes would be coded as recurrent).

Clinical patient-reported outcome measures included the change in patient-reported outcome measures for pain and hand function between baseline and 3 months, and health-related quality of life at 3 months. We used a VAS score (range 0 to 100) to measure pain during physical load (higher scores indicate more pain) and hand function (higher scores indicate better function). The VAS is a validated and widely used tool for measuring these constructs [38]. Although we also used more disease-specific questionnaires in our cohort (such as the Boston Carpal Tunnel Questionnaire, Patient-Rated Wrist/Hand Evaluation, and Michigan Hand outcomes Questionnaire), these differed among the treatments in our study sample and therefore are less well-suited to use for the current research question aiming at all patients with the most common hand and wrist conditions.

We measured health-related quality of life using the VAS of the EuroQol-5 Dimensions self-rated health questionnaire as an indication of the overall perceived health status (range 0 to 100; higher scores indicate better perceived health) [39].

To measure the patients’ experience with healthcare delivery (which is different from satisfaction with treatment results [40]), we used the PREM questionnaire, based on the Consumer Quality Index, which is widely used in private practice clinics in the Netherlands [41]. The 11 items evaluate the patients’ experience with healthcare delivery using a 5-point Likert scale (with answers ranging from no, not at all to yes, completely). Of this questionnaire, we only included three items because in the other items, ceiling effects were present that did not allow us to run our models. These items were experience with the explanation about the pros and cons of the treatment, experience with shared decision-making, and experience with the advice for at home.

Measures of mental health included anxiety and depression, pain catastrophizing, illness perceptions, and expectations. Anxiety and depression were measured with the Patient Health Questionnaire-4 (higher scores indicate more anxiety and depression) and pain catastrophizing was measured with the Pain Catastrophizing Scale (higher scores indicate a higher amount of catastrophizing). Illness perception was measured with the Brief Illness Perception Questionnaire [42,43]. The Brief Illness Perception Questionnaire measures how patients perceive their illness across eight domains (consequences, timeline, personal control, treatment control, identity, concern, coherence, and emotional response). Each domain is assessed with a single question (higher scores indicate more negative illness perceptions except for personal control, treatment control, and coherence) [44]. We excluded the domain of treatment control (“How much do you think your treatment

can help your illness?") from our analyses because we considered that item to have a strong conceptual overlap with the expectancy subscale of the Credibility/Expectancy questionnaire. Treatment outcome expectations were measured with the Credibility/Expectancy questionnaire [45]. The credibility subscale consists of three items measuring the credibility that the patient attributes to the treatment. A higher score reflects a higher attribution of credibility to a treatment. The expectancy subscale consists of three items measuring the expected magnitude of improvement because of the prescribed treatment. A higher score reflects a more positive treatment outcome expectation.

Ethical Approval

We obtained ethical approval for this study from Erasmus MC, Rotterdam, the Netherlands (MEC-2018-1088). Written informed consent was obtained from all patients.

Statistical Methods and Study Size

We dichotomized our outcome of satisfaction with treatment results into poor, moderate and fair as less satisfied, and good and excellent as more satisfied. After dichotomization, 57% (1042 of 1824) of participants were classified as more satisfied with the treatment results (19% [349 of 1824] answered excellent and 38% [693 of 1824] answered good), and 43% (782 of 1824) of patients were classified as less satisfied with the treatment results (26% [472 of 1824] reported their satisfaction was fair, 13% [231 of 1824] reported that it was moderate, and 4% [79 of 1824] reported that it was poor) (Fig. 2). This is comparable with other findings in our population [27,46-49]. Similarly, to further account for ceiling effects, we dichotomized the PREM items into negative experience (answering options: no, not at all, a little bit, partly, and mostly) and positive experience (answering option: yes, completely). The items used in the final analysis were: "Did you decide together with the care providers which care or treatment you will receive?" (hereinafter referred to as shared decision-making), "Have you been explained the pros and cons of the treatment or surgery?" (henceforth referred to as pros and cons), and "Were you advised by the healthcare providers on how to deal with your illness or complaints in your home situation?" (hereafter referred to as advice).

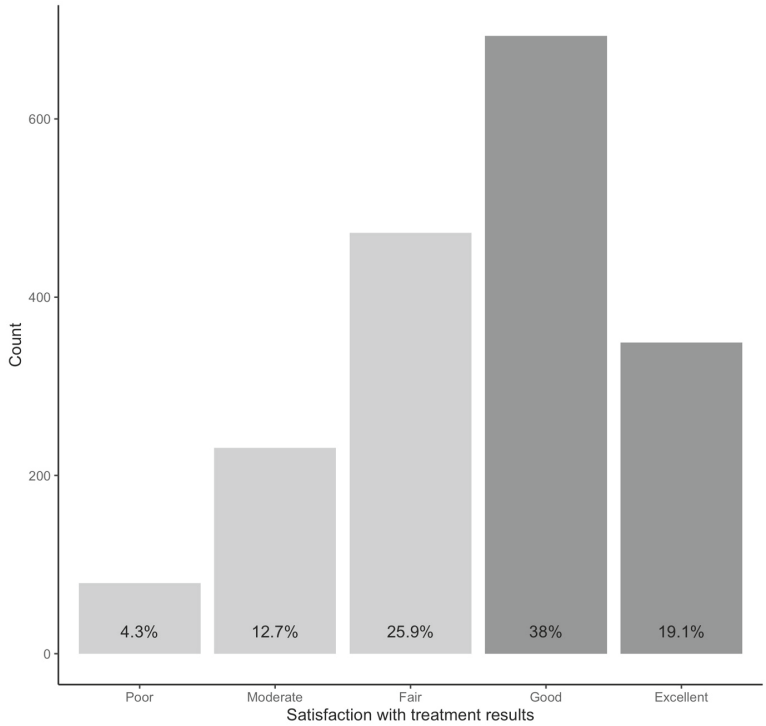


Fig. 2 This graph shows the distribution of satisfaction with treatment results at 3 months, before and after dichotomization. Light grey indicates patients who are less satisfied; dark grey indicates those who are more satisfied.

Because this study evaluated a diverse population of patients with common hand and wrist conditions, we adjusted for the type of treatment in the analyses. By adjusting for treatment in our analysis, we accounted for a potential influence of treatment on satisfaction with treatment results. To test the association of specific patient characteristics with satisfaction, we performed a hierarchical logistic regression analysis. In this hierarchical regression analysis, a set of variables is entered in a specific sequence to illustrate the added amount of explained variance of each set. In the first model, sociodemographic patient characteristics were entered, including age, sex, symptom duration, treatment side, dominance, type of work, and second-opinion visit. In the second step, we added clinical patient-reported outcome measures, including the EuroQol-5 Dimensions VAS self-rated health at 3 months, the change in VAS pain score during physical load, and VAS function score between baseline and 3 months. In the third step, we added the three items of the PREM: shared decision-making, pros and cons, and advice. In the fourth step, we added measures of mental health, including the Brief Illness Perception Questionnaire items of consequences, timeline, personal control, identity, concern, coherence, emotional response, Patient Health Questionnaire anxiety and depression subscales, Pain Catastrophizing Scale, and Credibility/Expectancy Questionnaire

subscales. An advantage of a hierarchical multivariable model is that by entering the next set of variables, certain variables might be pushed out of significance because variables may have shared variance. Therefore, in the most definitive multivariable model, only the variables that truly explain variance in the dependent variable remain. To account for potential strong correlations and multiple variables measuring the same construct, we evaluated multicollinearity using a correlation matrix (Supplementary Table 2; supplemental materials are available with the online version of *CORR*®) and variance inflation factor (Supplementary Table 3; supplemental materials are available with the online version of *CORR*®). A correlation coefficient of the Spearman rho greater than 0.7 was considered a strong correlation. A variance inflation factor greater than 3 was considered an indication of multicollinearity [50]. Based on the variance inflation factor (the highest variance inflation factor = 2.2) and the correlation matrix (highest Spearman rho = 6.8, which is only a moderate correlation), we did not find any indication for multicollinearity in the hierarchical logistic regression model. To illustrate the goodness of fit of the different models, we determined the area under the curve, the Nagelkerke r^2, and the receiver operating characteristic curves for each model.

With 1824 patients, 33 variables, an alpha of 0.05, and a conventional small effect size f^2 of 0.02, this study had a power of 95%. We additionally computed univariable associations between all variables. In addition to odds ratios, we reported standardized ORs by converting them to the same scale [51]. The nonstandardized odds ratios in our most definitive model indicate that with every unit increase in either a continuous, dichotomous, or categorical independent variable, the odds of being satisfied with the treatment results or being willing to undergo the treatment again increase or decrease by the value of the nonstandardized OR. Standardized ORs were converted to the same scale, which made it easier to make between-variable comparisons and determine the relative association of each explanatory variable.

All analyses were performed using R Statistical Programming, version 3.3.4 (R Project for Statistical Computing). For all tests, a p value < 0.05 was considered statistically significant.

Table 2. Most definitive model after the hierarchical logistic regression analyses (n = 1824) using sociodemographic, clinical characteristics, experience, and mental health characteristics explaining satisfaction with treatment results

Variables	Range (when applicable)	Nonstandardized OR (95% CI)	Standardized OR (95% CI)	p value
Age in years		0.99 (0.98-1.01)	0.92 (0.79-1.07)	0.27
Sex (male)		1.22 (0.95-1.59)	1.22 (0.95-1.59)	0.13
BMI		0.98 (0.96-1.00)	0.91 (0.81-1.02)	0.11
Dominant side treated (yes)		0.87 (0.69-1.10)	0.87 (0.69-1.10)	0.24
Workload (reference = unemployed)				
Light		1.04 (0.76-1.42)	1.04 (0.76-1.42)	0.81
Moderate		1.07 (0.77-1.48)	1.07 (0.77-1.48)	0.70
Severe		0.79 (0.50-1.24)	0.79 (0.50-1.24)	0.30
Symptom duration in months		1.00 (1.00-1.00)	1.03 (0.91-1.16)	0.66
Second opinion (no)		1.02 (0.48-2.18)	1.02 (0.48-2.18)	0.96
Recurrence (yes)		0.95 (0.63-1.45)	0.95 (0.63-1.45)	0.81
Smoking (no)		0.94 (0.67-1.32)	0.94 (0.67-1.32)	0.73
EQ-5D VAS self-rated health	0-100	1.01 (1.00-1.01)	1.13 (1.00-1.28)	0.05
Change in VAS pain during load	0-100	1.03 (1.02-1.03)	2.52 (2.18-2.92)	< 0.001
Change in VAS function	0-100	1.02 (1.01-1.02)	1.76 (1.54-2.01)	< 0.001
PREM shared decision-making positive (yes)		1.04 (0.80-1.36)	1.04 (0.80-1.36)	0.77
PREM pros/cons positive (yes)		1.83 (1.41-2.38)	1.83 (1.41-2.38)	< 0.001
PREM advice positive (yes)		1.57 (1.21-2.04)	1.57 (1.21-2.04)	< 0.001
B-IPQ consequences	0-10	0.95 (0.89-1.01)	0.88 (0.75-1.04)	0.12
B-IPQ timeline	0-10	1.06 (1.01-1.12)	1.20 (1.04-1.37)	0.01
B-IPQ personal control	0-10	1.09 (1.04-1.14)	1.24 (1.10-1.40)	< 0.001
B-IPQ identity	0-10	0.93 (0.88-0.99)	0.84 (0.72-0.97)	0.02
B-IPQ concern	0-10	0.94 (0.89-1.00)	0.84 (0.72-0.99)	0.04
B-IPQ coherence	0-10	0.98 (0.92-1.04)	0.95 (0.84-1.08)	0.43
B-IPQ emotional response	0-10	1.00 (0.95-1.06)	1.01 (0.86-1.18)	0.94
CEQ credibility score	3-27	1.03 (0.98-1.08)	1.11 (0.95-1.30)	0.19
CEQ expectancy score	3-27	1.05 (1.01-1.09)	1.23 (1.04-1.46)	0.02
PCS total score	0-52	0.99 (0.97-1.00)	0.90 (0.78-1.04)	0.17
PHQ-4 total score	0-12	1.01 (0.95-1.08)	1.03 (0.89-1.18)	0.70

Nonstandardized and standardized odds ratios, 95% CIs, and p values are displayed, along with the AUC and the Nagelkerke r^2 for the model; the nonstandardized odds ratios in our most definitive model indicate that with every unit increase in either a continuous, dichotomous, or categorical independent variable, the odds of being satisfied with the treatment results increase or decrease by the value of the nonstandardized OR; standardized odds ratio are converted to the same scale, which makes it easier to make between-variable comparisons and determine the relative association of each explanatory variable; interpretation AUC (ability of the model to discriminate between more satisfied and less satisfied patients) = 0.82; interpretation Nagelkerke r^2 (goodness of fit of the model) = 0.39; EQ5D = EuroQol-5 Dimensions; PREM = Patient-Reported Experience Measures; B-IPQ = Brief Illness Perception Questionnaire; CEQ = Credibility/Expectancy Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire; OR = Odds Ratio; SOR = Standardized Odds Ratio.

Results

Satisfaction with Treatment Results

In our most definitive model, we found an area under the curve of 0.82 (Table 2), indicating an excellent ability to distinguish more satisfied from less satisfied patients [52]. Satisfaction with the treatment results was associated with the following variables (arranged from the largest to the smallest standardized OR): greater decrease in pain during physical load (SOR 2.52 [95% CI 2.18 to 2.92]; p < 0.001), patient's positive experience with the explanation of the pros and cons of the treatment (determined with the question: "Have you been explained the pros and cons of the treatment or surgery?") (SOR 1.83 [95% CI 1.41 to 2.38]; p < 0.001), greater improvement in hand function (SOR 1.76 [95% CI 1.54 to 2.01]; p < 0.001), patients' positive experience with the advice for at home (determined with the question: "Were you advised by the healthcare providers on how to deal with your illness or complaints in your home situation?") (SOR 1.57 [95% CI 1.21 to 2.04]; p < 0.001), patient's better personal control (determined with the question: "How much control do you feel you have over your illness?") (SOR 1.24 [95% CI 1.10 to 1.40]; p < 0.001), patient's more positive treatment expectations (SOR 1.23 [95% CI 1.04 to 1.46]; p = 0.02), longer expected illness duration by the patient (SOR 1.20 [95% CI 1.04 to 1.37]; p = 0.01), a smaller number of symptoms the patient saw as part of the illness (SOR 0.84 [95% CI 0.72 to 0.97]; p = 0.02), and less concern about the illness the patient experiences (SOR 0.84 [95% CI 0.72 to 0.99]; p = 0.04) (Fig. 3). When analyzing the separate steps of the different models, sociodemographics alone provided an area under the curve (AUC) of 0.60 (95% CI 0.57 to 0.62) for the level of satisfaction with treatment results. When adding clinical characteristics, the AUC was 0.79 (95% CI 0.77 to 0.81). This further increased to 0.81 (95% CI 0.79 to 0.81) when adding PREMs, and finally, the AUC increased to 0.82 (95% CI 0.80 to 0.84) for the level of satisfaction with treatment results (Fig. 4).

Analyzing differences in variables between the different steps of the model for satisfaction with treatment results, we found that there were two differences (Supplementary Table 4; supplemental materials are available with the online version of *CORR*®). First, in Model 1, recurrence (determined with the question: "Have you been treated for the same disease before?") was associated with a smaller probability of being satisfied with the treatment results (standardized OR 0.70 [95% CI 0.50 to 1.00]), but after adding the clinical patient-reported outcome measures in Model 2, there was no association. This implies that a different change in patient-reported outcome measure score has a shared variance with recurrence and pushes recurrence out of significance. This means that a different change in patient-reported outcome measure score is the stronger variable. Second, a higher EuroQol-5 Dimensions self-rated health score was associated with a larger probability of being satisfied with the treatment results in Model 2 (standardized OR 1.32 [95% CI 1.18 to 1.48) and Model 3 (standardized OR 1.29 [95% CI 1.15 to 1.45]). However, after adding measures of mental health and treatment expectations in Model 4, we found that the EuroQol-5 Dimensions self-rated health score was no longer associated, and several

illness perception items and more positive expectations became associated with being satisfied with the treatment results. This finding suggests that EuroQol-5 Dimensions self-rated health has shared variance with specific measures of mental health, such as illness perception. This means that the mental health measures are the stronger variables (Supplementary Table 5; supplemental materials are available with the online version of *CORR*®).

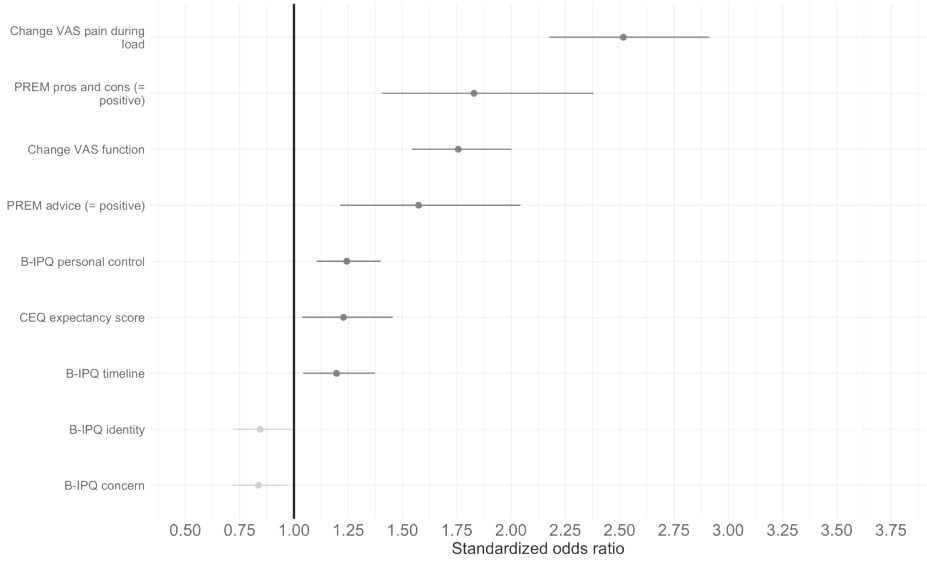


Fig. 3 This figure shows the standardized ORs of the associated variables for patient satisfaction with treatment results. Positive associations are shown in dark grey; negative associations are shown in light grey; PREM = patient-reported experience measures; CEQ = Credibility/Expectancy Questionnaire; B-IPQ = Brief Illness Perception Questionnaire.

Willingness to Undergo the Treatment Again

In our most definitive model, we found an area under the curve of 0.81 (Table 3), indicating an excellent ability to distinguish patients that would be willing to undergo the treatment again from patients that would not [52]. Being willing to undergo the treatment again was associated with the following variables (arranged from the largest to the smallest standardized OR): patient's positive experience with the information about the pros and cons (determined with the question: "Have you been explained the pros and cons of the treatment or surgery?") (SOR 2.05 [95% CI 1.50 to 2.8]; $p < 0.001$), greater improvement in hand function (SOR 1.80 [95% CI 1.54 to 2.11]; $p < 0.001$), greater decrease in pain during physical load (SOR 1.74 [95% CI 1.48 to 2.07]; $p < 0.001$), patient's positive experience with the advice for at home (determined with the question: "Were you advised by the healthcare providers on how to deal with your illness or complaints in your home situation?") (SOR 1.52 [95% CI 1.11 to 2.07]; $p = 0.01$), patient's positive experience with shared decision-making (determined with the question: "Did you decide together with

the care providers which care or treatment you will receive?") (SOR 1.45 [95% CI 1.06 to 1.99]; p = 0.02), higher credibility the patient attributes to the treatment (SOR 1.44 [95% CI 1.20 to 1.73]; p < 0.001), longer symptom duration (SOR 1.27 [95% CI 1.09 to 1.52]; p < 0.01), and patient's better understanding of the condition (SOR 1.17 [95% CI 1.01 to 1.34]; p = 0.03) (Fig. 5).

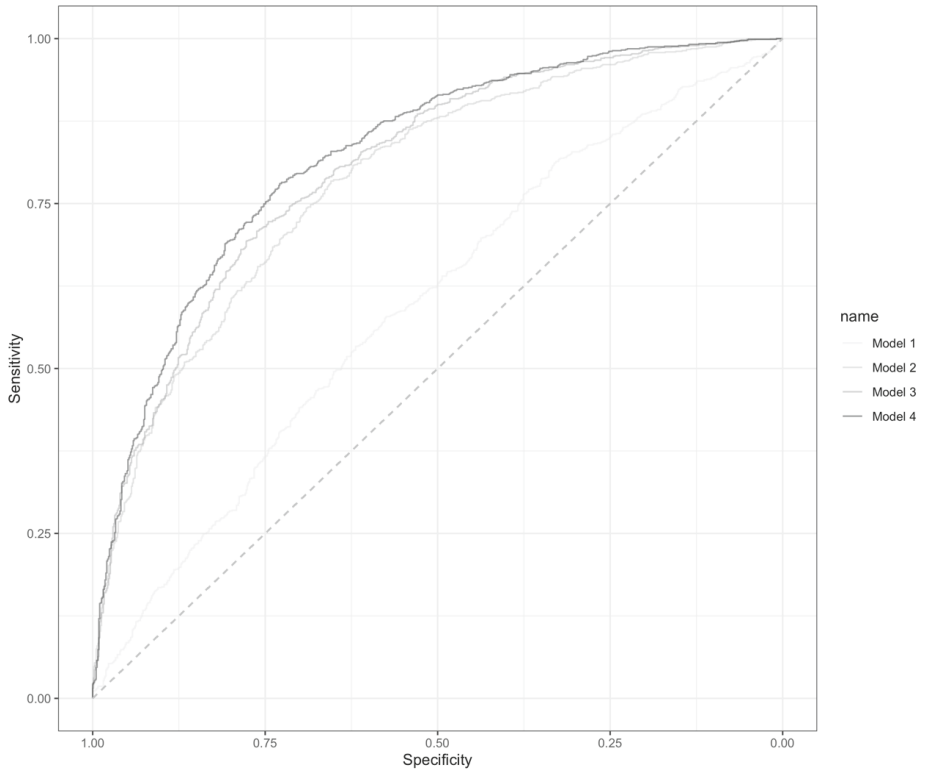


Fig. 4 This graph shows the area under the receiver operating characteristic curve for all models explaining the level of satisfaction with treatment results, using the 5-point Likert scale (question 1). The dashed line indicates a discriminative ability of 0.50. Model 1, including sociodemographics, had an AUC of 0.60, and Model 2, after adding clinical patient-reported outcome measures, had an AUC of 0.79. Model 3, after adding PREMs, had an AUC of 0.81, and after adding measures of mental health, the most definitive model had an AUC of 0.82.

For the willingness to undergo treatment again, sociodemographics alone provided an AUC of 0.58 (95% CI 0.55 to 0.62). When adding clinical characteristics, the AUC was 0.75 (95% CI 0.72 to 0.78). This further increased to 0.79 (95% CI 0.77 to 0.82) when adding PREMs, and finally, the AUC was 0.81 (95% CI 0.78 to 0.83) for the willingness to undergo treatment again after adding measures of mental health (Fig. 6).

Table 3. Most-definitive model after the hierarchical logistic regression analyses (n = 1824) using sociodemographic, clinical characteristics, experience, and mental health characteristics explaining undergo treatment again

Variables	Range (when applicable)	Nonstandardized OR (95% CI)	Standardized OR (95% CI)	p value
Age in years		0.99 (0.97-1.01)	0.90 (0.75-1.09)	0.28
Sex (male)		1.11 (0.80-1.54)	1.11 (0.80-1.54)	0.53
BMI		0.99 (0.96-1.02)	0.93 (0.81-1.07)	0.33
Dominant side treated (yes)		0.84 (0.63-1.11)	0.84 (0.63-1.11)	0.23
Workload (reference = unemployed)				
Light		1.30 (0.87-1.93)	1.30 (0.87-1.93)	0.20
Moderate		0.85 (0.56-1.27)	0.85 (0.56-1.27)	0.42
Severe		0.77 (0.44-1.35)	0.77 (0.44-1.35)	0.35
Symptom duration in months		1.01 (1.00-1.01)	1.27 (1.09-1.52)	< 0.01
Second opinion (no)		1.30 (0.52-3.00)	1.30 (0.52-3.00)	0.55
Recurrence (yes)		1.00 (0.62-1.64)	1.00 (0.62-1.64)	0.99
Smoking (no)		0.87 (0.56-1.32)	0.87 (0.56-1.32)	0.51
EQ-5D VAS self-rated health	0-100	1.00 (0.99-1.01)	0.96 (0.82-1.12)	0.65
Change in VAS pain during load	0-100	1.02 (1.01-1.02)	1.74 (1.48-2.07)	< 0.001
Change in VAS function	0-100	1.02 (1.01-1.02)	1.80 (1.54-2.11)	<0.001
PREM shared decision-making positive (yes)		1.45 (1.06-1.99)	1.45 (1.06-1.99)	0.02
PREM pros cons positive (yes)		2.05 (1.50-2.80)	2.05 (1.50-2.80)	< 0.001
PREM advice positive (yes)		1.52 (1.11-2.07)	1.52 (1.11-2.07)	0.01
B-IPQ consequences	0-10	0.95 (0.87-1.02)	0.87 (0.71-1.06)	0.17
B-IPQ timeline	0-10	1.01 (0.95-1.07)	1.02 (0.86-1.21)	0.82
B-IPQ personal control	0-10	1.02 (0.97-1.08)	1.06 (0.92-1.23)	0.41
B-IPQ identity	0-10	1.00 (0.93-1.07)	1.00 (0.82-1.20)	0.96
B-IPQ concern	0-10	0.99 (0.92-1.06)	0.96 (0.79-1.17)	0.71
B-IPQ coherence	0-10	1.08 (1.01-1.16)	1.17 (1.01-1.34)	0.03
B-IPQ emotional response	0-10	1.00 (0.94-1.07)	1.01 (0.83-1.23)	0.93

Table 3. *(continued)*

Variables	Range (when applicable)	Nonstandardized OR (95% CI)	Standardized OR (95% CI)	p value
CEQ credibility score	3-27	1.11 (1.06-1.18)	1.44 (1.20-1.73)	< 0.001
CEQ expectancy score	3-27	0.99 (0.94-1.04)	0.96 (0.78-1.18)	0.71
PCS total score	0-52	1.00 (0.98-1.02)	0.97 (0.82-1.15)	0.73
PHQ-4 total score	0-12	0.98 (0.90-1.06)	0.96 (0.81-1.13)	0.59

Nonstandardized and standardized odds ratios, 95% CIs, and p values are displayed, along with the AUC and Nagelkerke's r^2 for the model; the non-standardized odds ratios in our most-definitive model indicate that with every unit increase in either a continuous, dichotomous, or categorical independent variable, the odds of being willing to undergo the treatment again increase or decrease by the value of the nonstandardized OR; standardized odds ratios are converted to the same scale, which makes it easier to make between-variable comparisons and determine the relative association of each explanatory variable; interpretation AUC (ability of the model to discriminate between willing or not willing to undergo again) = 0.81; interpretation of the Nagelkerke r^2 (goodness of fit of the model) = 0.29; EQ-5D = EuroQol-5 Dimensions; PREM = Patient-Reported Experience Measures; B-IPQ = Brief Illness Perception Questionnaire; CEQ = Credibility/Expectancy Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire; OR = Odds Ratio; SOR = Standardized Odds Ratio.

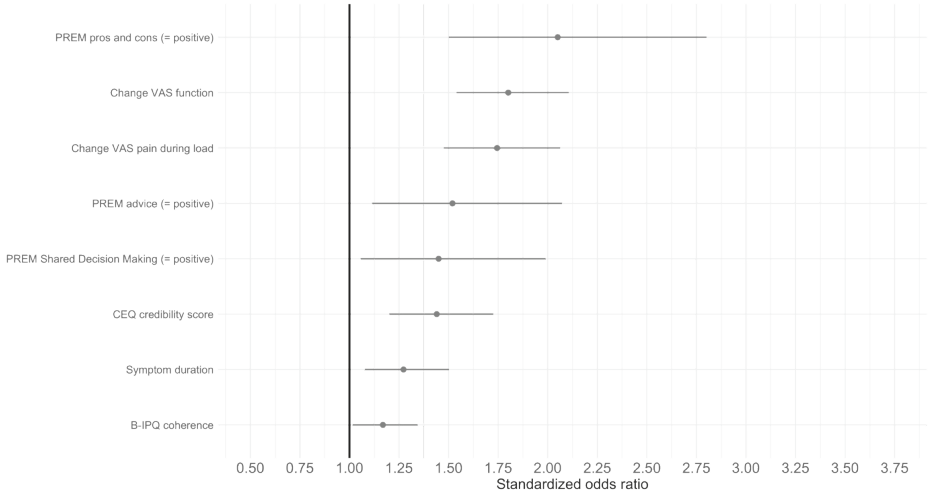


Fig. 5 This figure shows standardized ORs of the associated variables for the patient's willingness to undergo the treatment again; PREM = patient-reported experience measures; CEQ = Credibility/Expectancy Questionnaire; B-IPQ = Brief Illness Perception Questionnaire.

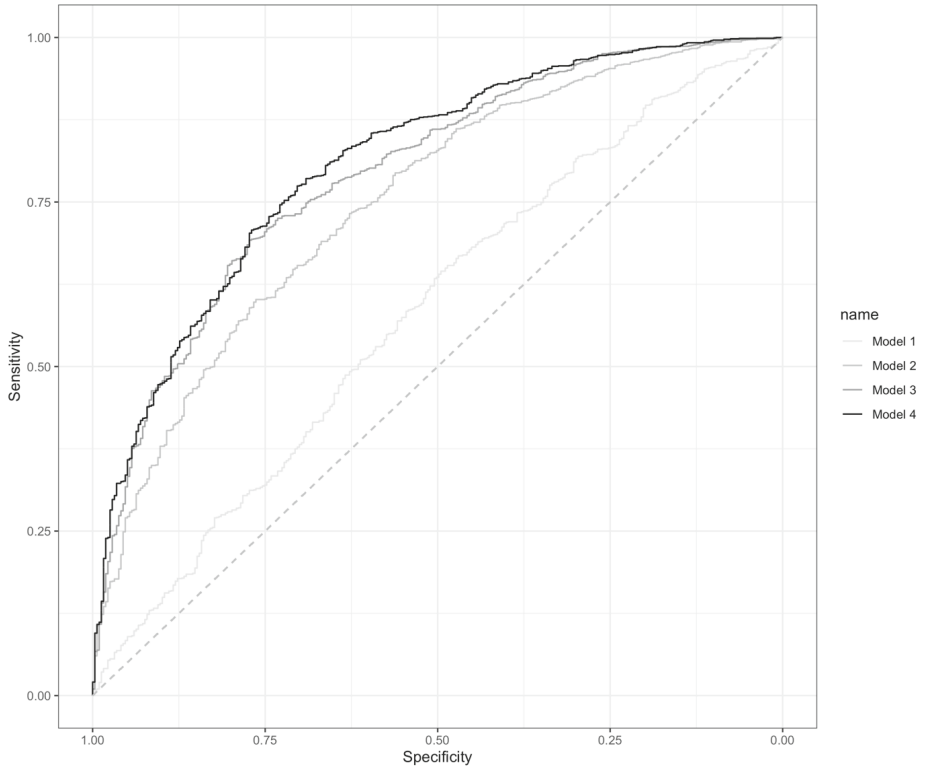


Fig. 6 This graph shows the area under the receiver operating characteristic curve for all models explaining the patient's willingness to undergo the treatment again (yes or no; question 2). The dashed line indicates a discriminative ability of 0.50. Model 1, including sociodemographics, had an AUC of 0.58, and after adding clinical patient-reported outcome measures, Model 2 had an AUC of 0.75. After adding PREMs, Model 3 had an AUC of 0.79. After adding measures of mental health, the most definitive model had an AUC of 0.81.

Discussion

In the framework of patient-centered and value-based healthcare, satisfaction with treatment results is an important outcome domain. Before our study, it was unclear which factors were independently associated with satisfaction with treatment results and with a willingness to undergo the treatment again. We found a high explained variance in our models. The following variables were independently associated with satisfaction in either or both models: greater decrease in pain during physical load, patient's positive experience with the explanation of the pros and cons of the treatment, positive experience with the advice for at home (determined with the question: "Were you advised by the healthcare providers on how to deal with your illness or complaints in your home situation?"), patient's positive experience with shared decision-making, higher credibility the patients attributes

to the treatment, longer symptom duration, better personal control (determined with the question: “How much control do you feel you have over your illness”), patient’s more positive treatment expectations, longer expected illness duration by the patient, patient’s better understanding of the condition, a smaller number of symptoms the patient sees as part of the illness, and less concern about the illness the patient experiences. Many of these variables may be guided and can be used directly in daily clinic or in studies that develop interventions to improve satisfaction with treatment results.

Limitations

Whereas an advantage of our observational study design is its representation of daily practice, a limitation of the observational design is that a substantial proportion of patients did not respond. However, the nonresponder analysis did not show substantial differences, and the Little test strongly suggests that the data were missing at random. Therefore, we are confident that the high percentage of nonresponders did not influence our results.

A second limitation is the follow-up time in our study. We chose this timepoint because follow-up measurements for the PREM were only obtained at 3 months. As a result, the more extensive surgical treatments may not have reached their endpoint yet, and evaluating satisfaction with treatment results may be too soon at this timepoint. However, theoretically, this should not influence factors explaining variance in satisfaction with treatment results. In fact, there might be more variation in satisfaction with treatment results at 3 months, which may yield better results. Nevertheless, future studies might investigate different timepoints.

Another limitation is the variety of treatment types in our study. Combining different treatment types may have led to dilution of the results because certain variables might interact with the treatment type. However, we aimed to investigate which factors explain satisfaction with treatment results in a general population of patients treated for hand and wrist disorders. Therefore, we selected the most commonly used treatment type in each of the six largest measurement tracks from our cohort and adjusted for the treatment type in our models. By adjusting for the treatment type in our analysis, a potential influence of treatment type on satisfaction with treatment results is accounted for, and the remaining significant variables are independent of treatment type in the final hierarchical model. Therefore, these remaining variables can be generalized to a broader population of patients with hand and wrist conditions. The standardized mean difference between the treatment types was small. This further strengthens the generalizability of our study findings, perhaps even to patients with other musculoskeletal conditions such as hip osteoarthritis. However, future studies should validate our findings in other populations.

Additionally, because satisfaction with treatment results is a multidimensional construct, there are still doubts about the validity of instruments measuring this domain [8,53]. Although the Satisfaction with Treatment Result Questionnaire has a good-to-excellent construct

validity and a very high test-retest reliability, future studies should further investigate its face validity.

Finally, we found a very high proportion of the finding explained by the variables in our model. An explanation for the little unexplained variance may be that we did not include all relevant variables in our models, such as additional aspects of experiences with healthcare delivery, coping strategies, goal attainment, the occurrence of complications, personal injury lawsuits, social health, or the specific course of rehabilitation. Additionally, our dichotomization may be a reason for unexplained variance, although this also has added value because our model thereby distinguishes between more satisfied and less satisfied patients. Moreover, although the Brief Illness Perception Questionnaire and Patient Health Questionnaire are valid tools, they might be interpreted differently by individuals, and they function as screening tools and lack the conceptual depth of more extensive questionnaires. Because satisfaction with treatment results is a complex domain, using more comprehensive measures of mental health may yield an even larger proportion of explained variance. Future studies might include these variables when investigating satisfaction with treatment results.

Discussion of Key Findings

Interestingly, all three included PREM items (positive experience with the explanation of the pros and cons; advice for how to deal with the complaints at home; and shared decision making) were associated with one or both of the Satisfaction with Treatment Result Questionnaire questions (which were: Are you satisfied with the treatment result so far? And, would you be willing to undergo the treatment again under similar circumstances?). These findings confirm that the patients' experience with healthcare delivery is associated with their satisfaction with the result. Based on these findings, healthcare providers may try to improve the experience with healthcare delivery, that is, by always explaining the pros and cons of a treatment and by providing adequate advice on how to deal with the complaints at home (such as by sending e-mails with treatment-specific information and educational movies). Also, healthcare providers may strive for better shared-decision making. Future research should inform if this will indeed improve satisfaction with treatment results.

In contrast to previous studies [8,20,23,53,54], depression was not associated with satisfaction in our most-definitive model. However, we did find a univariable association. This suggests that depression has a shared variance with other variables in our models; for example, other mental health items, such as the Illness Perception Questionnaire item of emotional response. Similarly, we did not find an association with pain catastrophizing, while other studies did [20-22,55]. No other study on this topic that we know of has investigated the association of depression or pain catastrophizing in combination with illness perception, which may explain why our findings are different from those reported by others.

Another interesting finding here was that a higher score on the Credibility/Expectancy Questionnaire expectancy subscale (the more positive expectations a patient has of a treatment) was associated with better satisfaction with treatment results. This is especially noteworthy because several studies have suggested that clinicians ought to try to work to temper patients' expectations [24-26,56], and many surgeons believe that it is important to help patients to cultivate reasonable expectations before surgery. By contrast, several other studies have suggested that boosting expectations is associated with better outcomes [13,27-30,57]. Our findings support the latter suggestion. Related to this, the credibility subscale (the extent to which a patient attributes credibility to a treatment) was associated with the patient's willingness to undergo the treatment again. To our knowledge, no other studies have investigated factors explaining this willingness to undergo treatment again, but it seems sensible that someone who does not find a treatment credible may be less willing to undergo that treatment again. Hence, it might be helpful to investigate possible interventions to boost expectations and improve the credibility of specific treatments.

A possible intervention to influence the experience with healthcare delivery, expectations, and illness perception may, for example, be the creation of a decision-support tool to specifically investigate the patients' needs for the clinician to respond accordingly. Further, future studies should investigate whether satisfaction with treatment results can be predicted (instead of explained, such as in this study), so that a prediction model could be used as a decision tool and to show what outcomes the patient may expect. Another option is to provide more personalized information relevant to the patient, such as emailing treatment-specific pros and cons. Additionally, to influence illness perception, future studies might investigate the effect of discussing illness perceptions and expectations during the first consultation. However, these suggestions are all hypothetical and future research should investigate their added value.

Conclusions

We identified several influenceable factors independently associated with satisfaction with treatment results. To directly improve satisfaction with treatment results, clinicians might seek to: (1) improve the patient's experience with healthcare delivery, (2) try to influence illness perception, and (3) boost treatment expectations and credibility. However, these recommendations are all hypothetical, and future research should investigate their added value. Moreover, future studies should investigate whether satisfaction with treatment results can be predicted (instead of explained, as was done in this study), so that a prediction model could be used as a decision-support tool that may inform shared-decision making and expectation management. Also, decision-support tools that investigate patient-specific needs may positively influence experience with healthcare delivery, expectations, and illness perception, which in turn may improve satisfaction with treatment results.

Acknowledgments

We thank all the patients who completed questionnaires as part of their clinical care and who agreed that their data could be used anonymously for the present study. We thank the caregivers, and other personnel of Xpert Clinics and Equipe Zorgbedrijven for assisting in the routine outcome measurements that are the basis for this study.

Group Authors

Members of the Hand-Wrist Study Group include: Dirk-Johannes Jacobus Cornelis van der Avoort MD, Richard Arjen Michiel Blomme MD, Herman Luitzen de Boer MD, Gijs Marijn van Couwelaar MD, Jan Debeij MD PhD, Jak Dekker MSc, Alexandra Fink PT, Klazina Paulina de Haas MD, Kennard Harmsen MD, Steven Eric Ruden Hovius MD PhD, Rob van Huis PT, Richard Koch MD, Alexander Kroeze MD, Thybout Matthias Moojen MD PhD, Mark Johannes Willem van der Oest PhD, Pierre-Yves Alain Adriaan Pennehouat PT, Willemijn Anna de Ridder PT MSc, Johannes Pieter de Schipper MD, Karin Schoneveld PT MSc, Berbel Jeannee Rinel Sluijter MD PhD, Jeronimus Maria Smit MD PhD, Xander Smit MD PhD, John Sebastiaan Souer MD PhD, Marloes Hendrina Paulina ter Stege MSc, Johannes Frederikes Maria Temming MD, Joris Sebastiaan Teunissen BSc, Jeroen Hein van Uchelen MD PhD, Joris Jan Veltkamp PT, Guus Maarten Vermeulen MD PhD, Erik Taco Walbeehm MD PhD, Oliver Theodor Zöphel MD PhD, Jelle Michiel Zuidam.

References

1. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. *N Engl J Med.* 2012;366(9):780-781.
2. Basch E. Patient-Reported Outcomes - Harnessing Patients' Voices to Improve Clinical Care. *N Engl J Med.* 2017;376(2):105-108.
3. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477-2481.
4. Goldhahn J, Angst F, Simmen BR. What counts: outcome assessment after distal radius fractures in aged patients. *J Orthop Trauma.* 2008;22(8 Suppl):S126-130.
5. Karnezis IA, Fragkiadakis EG. Association between objective clinical variables and patient-rated disability of the wrist. *J Bone Joint Surg Br.* 2002;84(7):967-970.
6. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev.* 2013;70(4):351-379.
7. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, Hovius SER. A Standard Set for Outcome Measurement in Patients With Hand and Wrist Conditions: Consensus by the International Consortium for Health Outcomes Measurement Hand and Wrist Working Group. *J Hand Surg Am.* 2021;46(10):841-855.e847.
8. Ring D, Leopold SS. Editorial-Measuring Satisfaction: Can It Be Done? *Clin Orthop Relat Res.* 2015;473(10):3071-3073.
9. De Ridder WA, van Kooij YE, Vermeulen GM, Slijper HP, Selles RW, Wouters RM. Test-retest Reliability and Construct Validity of the Satisfaction with Treatment Result Questionnaire in Patients with Hand and Wrist Conditions: A Prospective Study. *Clin Orthop Relat Res.* 2021;479(9):2022-2032.
10. Marks M, Herren DB, Vliet Vlieland TPM, Simmen BR, Angst F, Goldhahn J. Determinants of patient satisfaction after orthopedic interventions to the hand: A review of the literature. *J Hand Ther.* 2011;24(4):303-312.
11. Tsehaie J, van der Oest MJW, Poelstra R, Selles RW, Feitz R, Slijper HP, Hovius SER, Porsius JT. Positive experience with treatment is associated with better surgical outcome in trapeziometacarpal osteoarthritis. *J Hand Surg Eur Vol.* 2019;44(7):714-721.
12. Mandl LA, Galvin DH, Bosch JP, George CC, Simmons BP, Axt TS, Fossel AH, Katz JN. Metacarpophalangeal arthroplasty in rheumatoid arthritis: What determines satisfaction with surgery? *J Rheumatol.* 2002;29(12):2488-2491.
13. Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet.* 2001;357(9258):757-762.
14. Parrish RC, Menendez ME, Mudgal CS, Jupiter JB, Chen NC, Ring D. Patient Satisfaction and its Relation to Perceived Visit Duration with a Hand Surgeon. *J Hand Surg (USA).* 2016;41(2):257-262e254.
15. Bot AG, Bossen JK, Herndon JH, Ruchelsman DE, Ring D, Vranceanu AM. Informed shared decision-making and patient satisfaction. *Psychosomatics.* 2014;55(6):586-594.
16. Bernstein J, Kupperman E, Kandel LA, Ahn J. Shared Decision Making, Fast and Slow: Implications for Informed Consent, Resource Utilization, and Patient Satisfaction in Orthopaedic Surgery. *J Am Acad Orthop Surg.* 2016;24(7):495-502.

17. Pollak KI, Alexander SC, Tulsky JA, Lyna P, Coffman CJ, Dolor RJ, Gulbrandsen P, Ostbye T. Physician empathy and listening: associations with patient satisfaction and autonomy. *J Am Board Fam Med.* 2011;24(6):665-672.

18. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Health Prof.* 2004;27(3):237-251.

19. Menendez ME, Chen NC, Mudgal CS, Jupiter JB, Ring D. Physician Empathy as a Driver of Hand Surgery Patient Satisfaction. *J Hand Surg (USA).* 2015;40(9):1860-1865.

20. Vranceanu AM, Ring D. Factors associated with patient satisfaction. *J Hand Surg (USA).* 2011;36(9):1504-1508.

21. Mosegaard SB, Stilling M, Hansen TB. Higher preoperative pain catastrophizing increases the risk of low patient reported satisfaction after carpal tunnel release: a prospective study. *BMC Musculoskelet Disord.* 2020;21(1):42.

22. Mosegaard SB, Stilling M, Hansen TB. Pain Catastrophizing Scale as a predictor of low postoperative satisfaction after hand surgery. *J Orthop.* 2020;21:245-248.

23. Lozano Calderón SA, Paiva A, Ring D. Patient Satisfaction After Open Carpal Tunnel Release Correlates With Depression. *J Hand Surg (USA).* 2008;33(3):303-307.

24. Marks M, Audigé L, Reissner L, Herren DB, Schindele S, Vliet Vlieland TPM. Determinants of patient satisfaction after surgery or corticosteroid injection for trapeziometacarpal osteoarthritis: Results of a prospective cohort study. *Arch Orthop Trauma Surg.* 2015;135(1):141-147.

25. Kadzielski J, Malhotra LR, Zurakowski D, Lee SGP, Jupiter JB, Ring D. Evaluation of Preoperative Expectations and Patient Satisfaction After Carpal Tunnel Release. *J Hand Surg (USA).* 2008;33(10):1783-1788.

26. Frouzakis R, Herren DB, Marks M. Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis. *J Hand Surg Am.* 2015;40(3):483-490.

27. Hoogendam L, van der Oest MJW, Wouters RM, Andrinopoulou ER, Vermeulen GM, Slijper HP, Porsius JT, Selles RW. Patients With Higher Treatment Outcome Expectations Are More Satisfied With the Results of Nonoperative Treatment for Thumb Base Osteoarthritis: A Cohort Study. *Arch Phys Med Rehabil.* 2021;102(8):1533-1540.

28. Mandl LA, Burke FD, Shaw Wilgis EF, Lyman S, Katz JN, Chung KC. Could preoperative preferences and expectations influence surgical decision making? Rheumatoid arthritis patients contemplating metacarpophalangeal joint arthroplasty. *Plast Reconstr Surg.* 2008;121(1):175-180.

29. Swarup I, Henn CM, Gulotta LV, Henn RF, 3rd. Patient expectations and satisfaction in orthopaedic surgery: A review of the literature. *J Clin Orthop Trauma.* 2019;10(4):755-760.

30. Poelstra R, Selles RW, Slijper HP, van der Oest MJW, Feitz R, Hovius SER, Porsius JT. Better patients' treatment experiences are associated with better postoperative results in Dupuytren's disease. *J Hand Surg Eur Vol.* 2018;43(8):848-854.

31. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453-1457.

32. Selles RW, Wouters RM, Poelstra R, van der Oest MJW, Porsius JT, Hovius SER, Moojen TM, van Kooij Y, Pennehouat PY, van Huis R, Vermeulen GM, Feitz R, Slijper HP. Routine Health Outcome Measurement: Development, Design, and Implementation of the Hand and Wrist Cohort. *Plast Reconstr Surg.* 2020;146(2):343-354.

33. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. *P T.* 2008;33(12):700-711.

34. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.

35. de Groot JA, Janssen KJ, Zwinderman AH, Bossuyt PM, Reitsma JB, Moons KG. Correcting for partial verification bias: a comparison of methods. *Ann Epidemiol.* 2011;21(2):139-148.

36. Little RJA. A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association.* 1988;83(404):1198-1202.

37. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, Frangakis C, Hogan JW, Molenberghs G, Murphy SA, Neaton JD, Rotnitzky A, Scharfstein D, Shih WJ, Siegel JP, Stern H. The prevention and treatment of missing data in clinical trials. *N Engl J Med.* 2012;367(14):1355-1360.

38. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken).* 2011;63 Suppl 11:S240-252.

39. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L Better Than EQ-5D-3L? A Head-to-Head Comparison of Descriptive Systems and Value Sets from Seven Countries. *Pharmacoeconomics.* 2018;36(6):675-697.

40. York Health Economics Consortium. Patient-Reported Experience Measure (PREM) [online],. https://www.yhec.co.uk/glossary/patient-reported-experience-measure-prem/. Published 2016. Accessed November 30, 2021.

41. Delnoij DM, Rademakers JJ, Groenewegen PP. The Dutch consumer quality index: an example of stakeholder involvement in indicator development. *BMC Health Serv Res.* 2010;10:88.

42. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res.* 2006;60(6):631-637.

43. de Raaij EJ, Schröder C, Maissan FJ, Pool JJ, Wittink H. Cross-cultural adaptation and measurement properties of the Brief Illness Perception Questionnaire-Dutch Language Version. *Man Ther.* 2012;17(4):330-335.

44. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* 2009;50(6):613-621.

45. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry.* 2000;31(2):73-86.

46. Wouters RM, Slijper HP, Esteban Lopez L, Hovius SER, Selles RW. Beneficial Effects of Nonsurgical Treatment for Symptomatic Thumb Carpometacarpal Instability in Clinical Practice: A Cohort Study. *Arch Phys Med Rehabil.* 2020;101(3):434-441.

47. Wouters RM, Tsehaie J, Slijper HP, Hovius SER, Feitz R, Blomme A, Sluijter B, Schouten C, van der Avoort DJ, Walbeehm E, van Couwelaar G, Vermeulen G, Schipper H, Temming H, van Uchelen J, de Boer L, de Haas N, Zöphel O, Souer S, Hovius S, Moojen T, Smit X, van Huis R, Pennehouat PY, Schoneveld K, van Kooij Y, Wouters R, Zagt P, van Ewijk F, Moussault F, van Houwelingen R, Veltkamp J, Velde AT, Fink A, Slijper H, Selles R, Porsius J, Spekreijse K, Zhou C, Poelstra R, Janssen M, van der Oest M, Evers S, Dekker J, de Jong M, van Gestel J, Stege MT, Dekker M, Faber R, Santegoets F, Sieber-Rasch M, Gerritsen T, Selles RW. Exercise Therapy in Addition to an Orthosis Reduces Pain More Than an Orthosis Alone in Patients With Thumb Base Osteoarthritis: A Propensity Score Matching Study. *Arch Phys Med Rehabil.* 2019;100(6):1050-1060.

48. Tsehaie J, Spekreijse KR, Wouters RM, Slijper HP, Feitz R, Hovius SER, Selles RW. Outcome of a Hand Orthosis and Hand Therapy for Carpometacarpal Osteoarthritis in Daily Practice: A Prospective Cohort Study. *J Hand Surg Am.* 2018;43(11):1000-1009.e1001.

49. Zhou C, Hovius SER, Pieters AJ, Slijper HP, Feitz R, Selles RW. Comparative Effectiveness of Needle Aponeurotomy and Collagenase Injection for Dupuytren's Contracture: A Multicenter Study. *Plast Reconstr Surg Glob Open.* 2017;5(9):e1425.

50. Kutner MHN, C.J., Neter, J., Li, W. *Applied linear statistical models.* Boston: McGraw-Hill; 2005.

51. Sjölander A. Regression standardization with the R package stdReg. *Eur J Epidemiol.* 2016;31(6):563-574.

52. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied logistic regression.* Vol 398: John Wiley & Sons; 2013.

53. Sraj SA. CORR Insights®: Test-retest Reliability and Construct Validity of the Satisfaction with Treatment Result Questionnaire in Patients with Hand and Wrist Conditions: A Prospective Study. *Clin Orthop Relat Res.* 2021;479(9):2033-2035.

54. Ring D, Kadzielski J, Fabian L, Zurakowski D, Malhotra LR, Jupiter JB. Self-reported upper extremity health status correlates with depression. *J Bone Joint Surg Am.* 2006;88(9):1983-1988.

55. Ring D, Kadzielski J, Malhotra L, Lee SG, Jupiter JB. Psychological factors associated with idiopathic arm pain. *J Bone Joint Surg Am.* 2005;87(2):374-380.

56. Waljee JF, Chung KC. Commentary regarding "Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis". *J Hand Surg Am.* 2015;40(3):491-492.

57. Peerdeman KJ, van Laarhoven AIM, Keij SM, Vase L, Rovers MM, Peters ML, Evers AWM. Relieving patients' pain with expectation interventions: a meta-analysis. *Pain.* 2016;157(6):1179-1191.

Supplementary Table 1. Comparison of characteristics for patients who completed all questionnaires of interest at baseline and at 3 months (responders) and patients who did not complete all questionnaires of interest at 3 months or did not provide a relevant answer to a patient-reported experience measure item (non-responders)

Variable	Standardized mean difference
Age in years	0.15
Sex (men and women)	0.04
Treatment (all six treatments)	0.26
Second opinion (yes or no)	0.00
Hand dominance (right, left, or both)	0.02
Dominant hand treated (right, left, or both)	0.02
Symptom duration in months	0.05
Workload (not employed, light load, moderate load, severe load)	0.04
Pain Catastrophizing Scale score	0.05
Patient Health Questionnaire Score	0.05
CEQ Credibility Score	0.15
CEQ Expectancy Score	0.16
Brief Illness Perception Questionnaire Score	0.04

Supplementary Table 2. Correlation matrix showing the Spearman rho between all variables, including the excluded variable of B-IPQ treatment control

	Treatment	Age	Sex	BMI	Dominant side treated	Workload	Symptom duration	Second opinion	Smoking	Recurrence	EQ-5D self-rated health	Change VAS pain during load	Change VAS function
Treatment	1	0.13	0.17	-0.05	-0.02	-0.13	-0.08	0.07	-0.06	0.09	0.09	-0.07	-0.04
Age	0.13	1	0.19	-0.06	-0.06	-0.56	0.08	0.04	-0.19	0.01	0.08	-0.09	0
Gender	0.17	0.19	1	0.02	-0.03	-0.02	-0.02	0.04	-0.02	0.07	0.08	-0.13	-0.03
BMI	-0.05	-0.06	0.02	1	-0.02	0.02	-0.07	-0.01	-0.03	0	-0.17	0.12	0.05
Dominant side treated	-0.02	-0.06	-0.03	-0.02	1	0.03	0.01	-0.03	0.01	0.04	0.03	0.04	0.07
Workload	-0.13	-0.56	-0.02	0.02	0.03	1	-0.08	-0.04	0.12	0	0.01	0.09	-0.01
Symptom duration	-0.08	0.08	-0.02	-0.07	0.01	-0.08	1	-0.04	0.01	0.06	0.02	-0.09	0.01
Second opinion	0.07	0.04	0.04	-0.01	-0.03	-0.04	-0.04	1	-0.04	-0.04	0.07	-0.02	0
Smoking	-0.06	-0.19	-0.02	-0.03	0.01	0.12	0.01	-0.04	1	-0.02	-0.07	0.02	0.03
Recurrence	0.09	0.01	0.07	0	0.04	0	0.06	-0.04	-0.02	1	-0.03	-0.06	-0.04
EQ-5D self-rated health	0.09	0.08	0.08	-0.17	0.03	0.01	0.02	0.07	-0.07	-0.03	1	-0.03	-0.02
Change VAS pain during load	-0.07	-0.09	-0.13	0.12	0.04	0.09	-0.09	-0.02	0.02	-0.06	-0.03	1	0.45
Change VAS function	-0.04	0	-0.03	0.05	0.07	-0.01	0.01	0	0.03	-0.04	-0.02	0.45	1
PREM SDM	0	-0.02	-0.02	0.08	-0.01	0.01	-0.03	-0.04	0.02	0	0.02	0.08	0.09
PREM pros cons	-0.02	0	-0.02	0.08	-0.02	0.01	0.03	0.01	0.03	-0.04	0.06	0.16	0.19
PREM advice	-0.01	0.02	-0.04	0.08	0.01	0	-0.02	-0.03	0	-0.03	0.04	0.17	0.14
B-IPQ consequences	-0.23	-0.16	-0.20	0.05	0.06	0.15	-0.01	-0.07	0.1	-0.03	-0.22	0.23	0.17
B-IPQ timeline	0	0.05	-0.06	-0.08	-0.06	-0.07	0.18	-0.01	0.02	0.06	-0.10	-0.11	-0.10
B-IPQ personal control	0	0.05	-0.06	0.01	-0.03	-0.02	0	0	0	-0.03	0.05	-0.01	-0.04
B-IPQ treatment control	0.06	0	0.06	0.03	0.06	-0.01	-0.05	0.04	0.01	-0.02	0.16	0.13	0.14
B-IPQ identity	-0.22	-0.19	-0.23	0.09	0.03	0.15	-0.02	-0.06	0.08	-0.03	-0.26	0.24	0.16
B-IPQ concern	-0.18	-0.11	-0.13	0.02	0	0.11	0.03	-0.09	0.06	0.04	-0.22	0.08	0.04
B-IPQ coherence	0.06	0.09	0.05	-0.03	-0.04	-0.03	0.01	0.02	-0.01	-0.02	0.11	0.01	0
B-IPQ emotional response	-0.2	-0.14	-0.17	0.02	-0.02	0.09	-0.01	-0.05	0.09	0	-0.29	0.11	0.04
CEQ credibility score	0.06	0.03	0.06	0.06	0.05	0	-0.06	0.02	0.04	-0.03	0.17	0.14	0.15
CEQ expectancy score	0.06	0.02	0.07	0.05	0.09	0.01	-0.07	0.02	0.02	0	0.2	0.15	0.16
PCS score	-0.20	-0.07	-0.14	0.04	0.02	0.06	0	-0.01	0.07	-0.02	-0.26	0.14	0.04
PHQ score	-0.13	-0.17	-0.14	0.03	-0.02	0.04	-0.01	-0.01	0.13	-0.02	-0.31	0.05	0.02

EQ5D = EuroQol-5 Dimensions; PREM = Patient-Reported Experience Measures; B-IPQ = Brief Illness Perception Questionnaire; CEQ = Credibility/Expectancy Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire.

PREM SDM	PREM pros cons	PREM advice	B-IPQ consequences	B-IPQ timeline	B-IPQ personal control	B-IPQ treatment control	B-IPQ identity	B-IPQ concern	B-IPQ coherence	B-IPQ emotional response	CEQ credibility score	CEQ expectancy score	PCS score	PHQ Score
0	-0.02	-0.01	-0.23	0	0	0.06	-0.22	-0.18	0.06	-0.20	0.06	0.06	-0.20	-0.13
-0.02	0	0.02	-0.16	0.05	0.05	0	-0.19	-0.11	0.09	-0.14	0.03	0.02	-0.07	-0.17
-0.02	-0.02	-0.04	-0.2	-0.06	-0.06	0.06	-0.23	-0.13	0.05	-0.17	0.06	0.07	-0.14	-0.14
0.08	0.08	0.08	0.05	-0.08	0.01	0.03	0.09	0.02	-0.03	0.02	0.06	0.05	0.04	0.03
-0.01	-0.02	0.01	0.06	-0.06	-0.03	0.06	0.03	0	-0.04	-0.02	0.05	0.09	0.02	-0.02
0.01	0.01	0	0.15	-0.07	-0.02	-0.01	0.15	0.11	-0.03	0.09	0	0.01	0.06	0.04
-0.03	0.03	-0.02	-0.01	0.18	0	-0.05	-0.02	0.03	0.01	-0.01	-0.06	-0.07	0	-0.01
-0.04	0.01	-0.03	-0.07	-0.01	0	0.04	-0.06	-0.09	0.02	-0.05	0.02	0.02	-0.01	-0.01
0.02	0.03	0	0.1	0.02	0	0.01	0.08	0.06	-0.01	0.09	0.04	0.02	0.07	0.13
0	-0.04	-0.03	-0.03	0.06	-0.03	-0.02	-0.03	0.04	-0.02	0	-0.03	0	-0.02	-0.02
0.02	0.06	0.04	-0.22	-0.1	0.05	0.16	-0.26	-0.22	0.11	-0.29	0.17	0.2	-0.26	-0.31
0.08	0.16	0.17	0.23	-0.11	-0.01	0.13	0.24	0.08	0.01	0.11	0.14	0.15	0.14	0.05
0.09	0.19	0.14	0.17	-0.1	-0.04	0.14	0.16	0.04	0	0.04	0.15	0.16	0.04	0.02
1	0.4	0.36	0.03	-0.09	0.01	0.13	0.01	-0.04	0.10	-0.03	0.18	0.14	-0.01	-0.05
0.4	1	0.41	0	-0.07	0.03	0.05	-0.02	-0.04	0.08	-0.05	0.16	0.10	-0.05	-0.10
0.36	0.41	1	0.04	-0.09	0.06	0.10	0.02	-0.03	0.09	-0.04	0.17	0.10	-0.02	-0.03
0.03	0	0.04	1	0.13	-0.05	0.04	0.60	0.50	-0.03	0.52	0.01	0	0.39	0.31
-0.09	-0.07	-0.09	0.13	1	0.02	-0.35	0.16	0.34	-0.02	0.24	-0.30	-0.40	0.16	0.13
0.01	0.03	0.06	-0.05	0.02	1	-0.04	-0.06	-0.03	-0.02	-0.04	-0.05	-0.09	-0.06	-0.03
0.13	0.05	0.1	0.04	-0.35	-0.04	1	0.01	-0.21	0.29	-0.17	0.61	0.68	-0.11	-0.07
0.01	-0.02	0.02	0.6	0.16	-0.06	0.01	1	0.49	-0.03	0.45	-0.02	-0.05	0.37	0.25
-0.04	-0.04	-0.03	0.5	0.34	-0.03	-0.21	0.49	1	-0.16	0.59	-0.23	-0.24	0.45	0.34
0.1	0.08	0.09	-0.03	-0.02	-0.02	0.29	-0.03	-0.16	1	-0.18	0.3	0.22	-0.17	-0.12
-0.03	-0.05	-0.04	0.52	0.24	-0.04	-0.17	0.45	0.59	-0.18	1	-0.19	-0.2	0.49	0.47
0.18	0.16	0.17	0.01	-0.3	-0.05	0.61	-0.02	-0.23	0.3	-0.19	1	0.67	-0.16	-0.12
0.14	0.10	0.10	0	-0.40	-0.09	0.68	-0.05	-0.24	0.22	-0.20	0.67	1	-0.14	-0.10
-0.01	-0.05	-0.02	0.39	0.16	-0.06	-0.11	0.37	0.45	-0.17	0.49	-0.16	-0.14	1	0.40
-0.05	-0.10	-0.03	0.31	0.13	-0.03	-0.07	0.25	0.34	-0.12	0.47	-0.12	-0.10	0.40	1

Supplementary Table 3. Variance inflation factor of every included variable

Variable	Generalized Variance Inflation Factor
Age	1.8
Sex	1.3
BMI	1.1
Dominant treated hand	1.1
Type work	1.6
Symptom duration	1.1
Second opinion	1.0
Recurrence	1.1
Smoking	1.1
EQ-5D self-rated health	1.3
Change VAS pain during loading	1.3
Change VAS function	1.1
PREM shared decision-making	1.3
PREM pros/cons	1.3
PREM advice	1.3
B-IPQ consequences	2.0
B-IPQ timeline	1.5
B-IPQ personal control	1.1
B-IPQ identity	1.8
B-IPQ concern	2.0
B-IPQ coherence	1.1
B-IPQ emotional response	2.0
CEQ credibility score	1.9
CEQ expectancy score	2.2
PCS score	1.6
PHQ Score	1.5

EQ5D = EuroQol-5 Dimensions; PREM = Patient-Reported Experience Measures; B-IPQ = Brief Illness Perception Questionnaire; CEQ = Credibility/Expectancy Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire

Supplementary Table 4. Beta coefficients for hierarchical logistic regression models explaining satisfaction with treatment results 3 months after treatment

	Model 1		**Model 2**	
Explanatory variables	OR (95% CI)	SOR (95% CI)	OR (95% CI)	SOR (95% CI)
Age in years	1.00 (0.99-1.01)	1.02 (0.90-1.15)	1.00 (0.99-1.01)	0.98 (0.85-1.12)
Sex (male)	1.10 (0.89-1.36)	1.10 (0.89-1.36)	1.23 (0.96-1.57).	1.23 (0.96-1.57).
BMI	1.00 (0.98-1.02)	0.98 (0.89-1.08)	0.99 (0.97-1.01)	0.95 (0.85-1.06)
Dominant side treated (yes)	0.91 (0.75-1.11)	0.91 (0.75-1.11)	0.84 (0.67-1.04)	0.84 (0.67-1.04)
Workload (unemployed)				
Light	1.10 (0.85-1.42)	1.10 (0.85- 1.42)	0.98 (0.73-1.32)	0.98 (0.73-1.32)
Moderate	1.10 (0.83-1.45)	1.10 (0.83-1.45)	1 (0.73-1.36)	1 (0.73-1.36)
Severe	0.85 (0.58-1.24)	0.85 (0.58-1.24)	0.70 (0.46-1.08)	0.70 (0.46-1.08)
Symptom duration in months	1.00 (1.00-1.00)	0.99 (0.90-1.09)	1.00 (1.00-1.00)	1.04 (0.93-1.16)
Second opinion (no)	1.25 (0.66-2.33)	1.25 (0.66-2.33)	1.17 (0.58-2.38)	1.17 (0.58-2.38)
Recurrence (yes)	0.70 (0.50-1.0)[a]	0.70 (0.50-1.0)[a]	0.86 (0.58-1.28)	0.86 (0.58-1.28)
Smoking (no)	0.96 (0.73-1.28)	0.96 (0.73-1.28)	0.91 (0.66-1.26)	0.91 (0.66-1.26)
EQ-5D self-rated health			1.01 (1.01-1.02)[c]	1.32 (1.18-1.48)[c]
Change in VAS pain during load			1.03 (1.02-1.03)[c]	2.44 (2.13-2.81)[c]
Change in VAS function			1.02 (1.01-1.02)[c]	1.76 (1.56-1.99)[c]
PREM shared decision-making satisfied (yes)				
PREM pros cons satisfied (yes)				
PREM advice satisfied (yes)				
B-IPQ consequences				
B-IPQ timeline				
B-IPQ personal control				
B-IPQ identity				
B-IPQ concern				
B-IPQ coherence				

Model 3		Model 4		Univariable models	
OR (95% CI)	SOR (95% CI)	OR (95% CI)	SOR (95% CI)	OR (95% CI)	SOR (95% CI)
1.00 (0.98-1.01)	0.96 (0.83-1.11)	0.99 (0.98-1.01)	0.92 (0.79-1.07)	1.00 (0.90-1.00)	0.97 (0.89-1.10)
1.25 (0.97-1.61).	1.25 (0.97-1.61).	1.22 (0.95-1.59)	1.22 (0.95-1.59)	1.10 (0.92-1.30)	1.10 (0.92-1.30)
0.98 (0.96-1.01)	0.92 (0.82-1.03)	0.98 (0.96-1.00)	0.91 (0.81-1.02)	1.00 (0.98-1.00)	0.99 (0.91-1.10)
0.85 (0.68-1.06)	0.85 (0.68-1.06)	0.87 (0.69-1.10)	0.87 (0.69-1.10)	1.00 (0.85-1.20)	1.00 (0.85-1.20)
0.98 (0.73-1.33)	0.98 (0.73-1.33)	1.04 (0.76-1.42)	1.04 (0.76-1.42)	1.09 (0.87-1.40)	1.09 (0.87-1.40)
0.99 (0.72-1.36)	0.99 (0.72-1.36)	1.07 (0.77-1.48)	1.07 (0.77-1.48)	1.10 (0.87-1.40)	1.10 (0.87-1.40)
0.69 (0.45-1.07).	0.69 (0.45-1.07).	0.79 (0.50-1.24)	0.79 (0.50-1.24)	0.93 (0.67-1.30)	0.93 (0.67-1.30)
1 (1-1)	1.05 (0.93-1.18)	1 (1-1)	1.03 (0.91-1.16)	1.00 (1.00-1.00)	1.00 (0.91-1.10)
1.16 (0.56-2.41)	1.16 (0.56-2.41)	1.02 (0.48-2.18)	1.02 (0.48-2.18)	1.20 (0.65-2.20)	1.20 (0.65-2.20)
0.88 (0.59-1.33)	0.88 (0.59-1.33)	0.95 (0.63-1.45)	0.95 (0.63-1.45)	0.78 (0.55-1.10)	0.78 (0.55-1.10)
0.93 (0.67-1.3)	0.93 (0.67-1.3)	0.94 (0.67-1.32)	0.94 (0.67-1.32)	0.94 (0.71-1.20)	0.94 (0.71-1.20)
1.01 (1.01-1.02)[c]	1.29 (1.15-1.45)[c]	1.01 (1-1.01)	1.13 (1-1.28)	1.01 (1.01-1.02)[c]	1.30 (1.20-1.40)[c]
1.03 (1.02-1.03)[c]	2.37 (2.06-2.73)[c]	1.03 (1.02-1.03)[c]	2.52 (2.18-2.92)[c]	1.03 (1.03-1.03)[c]	2.50 (2.20-2.90)[c]
1.02 (1.01-1.02)[c]	1.68 (1.48-1.91)[c]	1.02 (1.01-1.02)[c]	1.76 (1.54-2.01)[c]	1.03 (1.02-1.03)[c]	2.30 (2.0-2.50)[c]
1.09 (0.84-1.41)	1.09 (0.84-1.41)	1.04 (0.8-1.36)	1.04 (0.8-1.36)	1.77 (1.45-2.10)[c]	1.77 (1.45-2.10)[c]
1.91 (1.47-2.46)[c]	1.91 (1.47-2.46)[c]	1.83 (1.41-2.38)[c]	1.83 (1.41-2.38)[c]	2.86 (2.35-3.50)[c]	2.86 (2.35-3.50)[c]
1.59 (1.23-2.04)[c]	1.59 (1.23-2.04)[c]	1.57 (1.21-2.04)[c]	1.57 (1.21-2.04)[c]	2.50 (2.05-3.05)[c]	2.50 (2.05-3.05)[c]
		0.95 (0.89-1.01)	0.88 (0.75-1.04)	0.98 (0.94-1.00)	0.94 (0.86-1.00)
		1.06 (1.01-1.12)[a]	1.2 (1.04-1.37)[a]	0.94 (0.91-0.97)[c]	0.84 (0.77-0.92)[c]
		1.09 (1.04-1.14)[c]	1.24 (1.1-1.4)[c]	1.1 (1.0-1.10)[a]	1.1 (1.0-1.20)[a]
		0.93 (0.88-0.99)[a]	0.84 (0.72-0.97)[a]	0.96 (0.93-0.99)[a]	0.90 (0.82-0.99)[a]
		0.94 (0.89-1)[a]	0.84 (0.72-0.99)[a]	0.92 (0.89-0.96)[c]	0.80 (0.73-0.88)[c]
		0.98 (0.92-1.04)	0.95 (0.84-1.08)	1.04 (0.99-1.10)	1.10 (0.98-1.20)

Supplementary Table 4. Beta coefficients for hierarchical logistic regression models explaining satisfaction with treatment results 3 months after treatment *(continued)*

	Model 1	**Model 2**
B-IPQ emotional response		
CEQ credibility score		
CEQ expectancy score		
PCS total score		
PHQ-4 total score		
AUC	0.60	0.79
Nagelkerke's r^2	0.04	0.32

In each additional model, more variables potentially explaining satisfaction with treatment results are included. Both the nonstandardized OR and standardized ORs are reported with 95% CIs. [a] $p \leq 0.05$; [b] $p \leq 0.01$; [c] $p \leq 0.001$.

The nonstandardized odds ratios indicate that with every unit increase in either a continuous, dichotomous, or categorical independent variable, the odds of being satisfied with the treatment results increase or decrease by the value of the nonstandardized OR; standardized odds ratio are converted to the same scale, which makes it easier to make between-variable comparisons and determine the relative association of each explanatory variable.

EQ5D = EuroQol-5 Dimensions; PREM = Patient-Reported Experience Measures; B-IPQ = Brief Illness Perception Questionnaire; CEQ = Credibility/Expectancy Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire; OR = Odds Ratio; SOR = Standardized Odds Ratio.

	Model 3	Model 4		Univariable models	
		1.00 (0.95-1.06)	1.01 (0.86-1.18)	0.95 (0.92-0.98)[c]	0.85 (0.77-0.93)[c]
		1.03 (0.98-1.08)	1.11 (0.95-1.3)	1.13 (1.10-1.16)[c]	1.50 (1.4-1.7)[c]
		1.05 (1.01-1.09)[a]	1.23 (1.04-1.46)[a]	1.11 (1.08-1.13)[c]	1.50 (1.4-1.7)[c]
		0.99 (0.97-1)	0.90 (0.78-1.04)	0.98 (0.97-0.99)[b]	0.86 (0.79-0.95)[b]
		1.01 (0.95-1.08)	1.03 (0.89-1.18)	0.95 (0.91-0.99)[a]	0.89 (0.81-0.98)[a]
	0.81	0.82		___	
	0.35	0.39		___	

Supplementary Table 5. Beta coefficients for hierarchical logistic regression models explaining willingness to undergo treatment again results 3 months after treatment

Variable	Model 1		Model 2	
Explanatory variables	OR (95% CI)	SOR (95% CI)	OR (95% CI)	SOR (95% CI)
Age in years	1.00 (0.98 - 1.01)	0.99 (0.84-1.16)	1.00 (0.98-1.01)	0.95 (0.80-1.13)
Sex (male)	1.00 (0.76-1.33)	1.00 (0.76-1.33)	1.08 (0.80-1.47)	1.08 (0.80-1.47)
BMI	1.00 (0.98-1.03)	1.01 (0.89-1.15)	1.00 (0.97-1.03)	0.98 (0.86-1.13)
Dominant side treated (yes)	0.87 (0.67-1.11)	0.87 (0.67-1.11)	0.80 (0.61-1.04).	0.80 (0.61-1.04).
Workload (unemployed)				
Light	1.29 (0.92-1.83)	1.29 (0.92-1.83)	1.25 (0.86-1.82)	1.25 (0.86-1.82)
Moderate	0.91 (0.64-1.30)	0.91 (0.64-1.30)	0.84 (0.57-1.23)	0.84 (0.57-1.23)
Severe	0.85 (0.53-1.38)	0.85 (0.53-1.38)	0.78 (0.47-1.31)	0.78 (0.47-1.31)
Symptom duration in months	1.00 (1.00-1.01)	1.14 (0.99-1.34)	1.01 (1.00-1.01)[a]	1.22 (1.05-1.44)[a]
Second opinion (no)	1.41 (0.62-2.90)	1.41 (0.62-2.90)	1.42 (0.6-3.07)	1.42 (0.6-3.07)
Recurrence (yes)	0.80 (0.53-1.24)	0.80 (0.53-1.24)	0.92 (0.59-1.47)	0.92 (0.59-1.47)
Smoking (no)	0.87 (0.59-1.25)	0.87 (0.59-1.25)	0.84 (0.56-1.25)	0.84 (0.56-1.25)
EQ5D self-rated health			1.01 (1.00-1.01)	1.12 (0.98-1.27)
Change in VAS pain during load			1.02 (1.01-1.02)[c]	1.81 (1.55-2.12)[c]
Change in VAS function			1.02 (1.01-1.02)[c]	1.84 (1.59-2.13)[c]
PREM Shared Decision Making Satisfied (yes)				
PREM Pros Cons Satisfied (yes)				
PREM Advice Satisfied (yes)				
B-IPQ consequences				
B-IPQ timeline				
B-IPQ personal control				
B-IPQ identity				
B-IPQ concern				
B-IPQ understanding				
B-IPQ emotional response				

Model 3		Model 4		Univariable models	
OR (95% CI)	SOR (95% CI)	OR (95% CI)	SOR (95% CI)	OR (95% CI)	SOR (95% CI)
0.99 (0.98-1.01)	0.94 (0.78-1.12)	0.99 (0.97-1.01)	0.90 (0.75-1.09)	0.99 (0.98-1.00)	0.94 (0.83-1.06)
1.14 (0.84-1.57)	1.14 (0.84-1.57)	1.11 (0.80-1.54)	1.11 (0.80-1.54)	0.88 (0.69-1.13)	0.88 (0.69-1.13)
0.99 (0.96-1.02)	0.94 (0.82-1.08)	0.99 (0.96-1.02)	0.93 (0.81-1.07)	1.01 (0.98-1.04)	1.04 (0.92-1.18)
0.82 (0.62-1.08)	0.82 (0.62-1.08)	0.84 (0.63-1.11)	0.84 (0.63-1.11)	0.88 (0.69-1.12)	0.88 (0.69-1.12)
1.28 (0.87-1.89)	1.28 (0.87-1.89)	1.30 (0.87-1.93)	1.30 (0.87-1.93)	1.27 (0.93-1.74)	1.27 (0.93-1.74)
0.84 (0.56-1.25)	0.84 (0.56-1.25)	0.85 (0.56-1.27)	0.85 (0.56-1.27)	0.97 (0.72-1.33)	0.97 (0.72-1.33)
0.76 (0.45-1.31)	0.76 (0.45-1.31)	0.77 (0.44-1.35)	0.77 (0.44-1.35)	0.96 (0.63-1.49)	0.96 (0.63-1.49)
1.01 (1.00-1.01)[b]	1.26 (1.08-1.49)[b]	1.01 (1.00-1.01)[b]	1.27 (1.09-1.52)[b]	1.00 (1.00-1.01)	1.07 (0.95-1.24)
1.48 (0.61-3.3)	1.48 (0.61-3.3)	1.30(0.52-3.00)	1.30 (0.52-3.00)	1.31 (0.58-2.64)	1.31 (0.58-2.64)
0.93 (0.58-1.51)	0.93 (0.58-1.51)	1.00 (0.62-1.64)	1.00 (0.62-1.64)	0.73 (0.49-1.12)	0.73 (0.49-1.12)
0.87 (0.56-1.30)	0.87 (0.56-1.30)	0.87 (0.56-1.32)	0.87 (0.56-1.32)	0.85 (0.58-1.21)	0.85 (0.58-1.21)
1.00 (1.00-1.01)	1.08 (0.94-1.24)	1.00 (0.99-1.01)	0.96 (0.82-1.12)	1.00 (1.00-1.01)	1.08 (0.96-1.21)
1.02 (1.01-1.02)[c]	1.72 (1.46-2.03)[c]	1.02 (1.01-1.02)[c]	1.74 (1.48-2.07)[c]	1.02 (1.02-1.03)[c]	2.09 (1.83-2.40)[c]
1.02 (1.01-1.02)[c]	1.76 (1.52-2.06)[c]	1.02 (1.01-1.02)[c]	1.80 (1.54-2.11)[c]	1.02 (1.02- 1.03)[c]	2.17 (1.90-2.48)[c]
1.59 (1.17-2.16)[b]	1.59 (1.17-2.16)[b]	1.45 (1.06-1.99)[a]	1.45 (1.06-1.99)[a]	2.50 (1.95-3.19)[c]	2.50 (1.95-3.19)[c]
2.17 (1.6-2.94)[c]	2.17 (1.6-2.94)[c]	2.05 (1.50-2.8)[c]	2.05 (1.5-2.8)[c]	3.89 (3.02-5.02)[c]	3.89 (3.02-5.02)[c]
1.57 (1.16-2.12)[b]	1.57 (1.16-2.12)[b]	1.52 (1.11-2.07)[b]	1.52 (1.11-2.07)[b]	2.91 (2.28-3.73)[c]	2.91 (2.28-3.73)[c]
		0.95 (0.87-1.02)	0.87 (0.71-1.06)	1.01 (0.97-1.06)	1.03 (0.92-1.17)
		1.01 (0.95-1.07)	1.02 (0.86-1.21)	0.95 (0.91-0.99)[a]	0.85 (0.76-0.97)[a]
		1.02 (0.97-1.08)	1.06 (0.92-1.23)	1.03 (0.98-1.08)	1.07 (0.95-1.20)
		1.00 (0.93-1.07)	1.00 (0.82-1.20)	1.01 (0.96-1.06)	1.03 (0.91-1.16)
		0.99 (0.92-1.06)	0.96 (0.79-1.17)	0.97 (0.92-1.01)	0.90 (0.80-1.02)
		1.08 (1.01-1.16)[a]	1.17 (1.01-1.34)[a]	1.13 (1.07-1.19)[c]	1.27 (1.13-1.41)[c]
		1.00 (0.94-1.07)	1.01 (0.83-1.23)	0.98 (0.94-1.02)	0.93 (0.82-1.05)

Supplementary Table 5. Beta coefficients for hierarchical logistic regression models explaining willingness to undergo treatment again results 3 months after treatment *(continued)*

Variable	Model 1	Model 2
CEQ credibility score		
CEQ expectancy score		
PCS total score		
PHQ-4 total score		
AUC	0.58	0.75
Nagelkerke's r^2	0.02	0.19

In each additional model, more variables potentially explaining satisfaction with treatment results are included. Both the nonstandardized OR and standardized ORs are reported with 95% CIs. [a] $p \leq 0.05$; [b] $p \leq 0.01$; [c] $p \leq 0.001$.

The nonstandardized odds ratios indicate that with every unit increase in either a continuous, dichotomous, or categorical independent variable, the odds of being satisfied with the treatment results increase or decrease by the value of the nonstandardized OR; standardized odds ratio are converted to the same scale, which makes it easier to make between-variable comparisons and determine the relative association of each explanatory variable.

EQ5D = EuroQol-5 Dimensions; PREM = Patient-Reported Experience Measures; B-IPQ = Brief Illness Perception Questionnaire; CEQ = Credibility/Expectancy Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire; OR = Odds Ratio; SOR = Standardized Odds Ratio.

	Model 3	Model 4		Univariable models	
		1.11 (1.06-1.18)[c]	1.44 (1.2-1.73)[c]	1.13 (1.10-1.17)[c]	1.52 (1.36-1.70)[c]
		0.99 (0.94-1.04)	0.96 (0.78-1.18)	1.06 (1.03-1.09)[c]	1.29 (1.15-1.44)[c]
		1.00 (0.98-1.02)	0.97 (0.82-1.15)	0.99 (0.98-1.00)	0.93 (0.82-1.04)
		0.98 (0.9-1.06)	0.96 (0.81-1.13)	0.95 (0.90-1.00)	0.90 (0.80-1.01)
	0.79	0.81			
	0.26	0.29			

PART 2

EXPLORE THE CONNECTION WITH THE PATIENT'S MINDSET

4

TREATMENT INVASIVENESS AND ILLNESS PERCEPTIONS ARE STRONGLY ASSOCIATED WITH OUTCOME EXPECTATIONS IN PATIENTS TREATED FOR HAND OR WRIST CONDITIONS: A CROSS-SECTIONAL STUDY

de Ridder WA, Hoogendam L, Zeroual F, Slijper HP, Wouters RM, Vermeulen GM, Selles RW, van der Oest MJW, the Hand-Wrist Study Group; 2023; Clinical Orthopaedics and Related Research

Abstract

Background

Multiple studies have shown that more-positive outcome expectations are associated with better treatment outcomes. Although this has not been shown to represent a causal relationship, there nonetheless is an interest in positively modifying outcome expectations to improve treatment outcomes. However, little is known about what is independently associated with outcome expectations in clinical practice. For example, it is unknown to what extent expectations are associated with contextual factors such as treatment or patient characteristics such as sociodemographics, or patient-reported outcome measures (PROMs) on patient perceptions of physical or mental health or illness. Studying factors associated with outcome expectations may provide relevant information for clinicians and researchers aiming to improve outcome expectations. Improving expectations might, in turn, improve treatment outcomes.

Question/purpose

Which factors (that is, sociodemographics, PROMs, illness perceptions, treatment, surgeon, and location) are independently associated with outcome expectations in patients with hand or wrist conditions?

Methods

This was a cross-sectional study. Between July 2018 and December 2021, we screened 21,327 patients with a diagnosed hand or wrist condition with complete baseline sociodemographic data such as age and workload. Sixty percent (12,765 of 21,327) of patients completed all relevant PROMs. We excluded patients receiving rare treatments, leaving 58% (12,345 of 21,327) for inclusion in the final sample. Those who participated were more often scheduled for surgical treatment and had higher expectations. We performed a multilevel analysis involving two steps. First, we evaluated whether patients receiving the same treatment, being counseled by the same surgeon, or being treated at the same location have more similar outcome expectations. We found that only patients receiving the same treatment had more similar outcome expectations. Therefore, we used a multilevel regression model to account for this correlation within treatments, and added treatment characteristics (such as nonsurgical versus minor or major surgery, which explained the effectiveness of each treatment) to potential explanatory factors. Second, in the multilevel hierarchical regression analysis, we added sociodemographics (Model 1), PROMs for physical and mental health (Model 2), illness perceptions (Model 3), and treatment characteristics (most-definitive model) to assess the explained variance in outcome expectations per step and the relative association with outcome expectations.

Results

Sociodemographic factors such as age and workload explained 1% of the variance in outcome expectations. An additional 2% was explained by baseline PROMs for physical

and mental health, 9% by illness perceptions, and 18% by treatment characteristics, resulting in an explained variance of 29% of the most-definitive model. A large number of patient and treatment characteristics were associated with outcome expectations. We used standardized betas to compare the magnitude of the effect of the different continuous and categorical variables. Among the associated variables, minor surgery (standardized beta [β] = 0.56 [95% confidence interval 0.44 to 0.68]; $p < 0.001$) and major surgery (β = 0.61 [95% CI 0.49 to 0.73]; $p < 0.001$) had the strongest positive association with outcome expectations (receiving surgery is associated with higher outcome expectations than nonsurgical treatment). A longer illness duration expected by the patient (-0.23 [95% CI -0.24 to -0.21]; $p < 0.001$) and being treated for the same condition as before (-0.08 [95% CI -0.14 to -0.03]; $p = 0.003$) had the strongest negative association with outcome expectations.

Conclusions

Outcome expectations are mainly associated with the invasiveness of the treatment and by patients' illness perceptions; patients before surgical treatment have more positive expectations of the treatment outcome than patients before nonsurgical treatment, even after accounting for differences in clinical and psychosocial profiles. In addition, patients with a more-positive perception of their illness had more-positive expectations of their treatment. Our findings suggest expectation management should be tailored to the specific treatment (such as surgical versus nonsurgical) and the specific patient (including their perception of their illness). It may be more beneficial to test and implement expectation management strategies for nonsurgical treatments such as physical therapy than for surgical treatments, given that our findings indicate a greater need to do so. An additional advantage of such a strategy is that successful interventions may prevent converting to surgical interventions, which is a goal of the stepped-care principles of standard care. Future studies might investigate the causality of the association between pretreatment expectations and outcomes by performing an experimental study such as a randomized controlled trial, in which boosting expectations is compared with usual care in nonsurgical and surgical groups.

Introduction

Patients have expectations at the beginning of their treatments regarding potential outcomes. Several studies have shown these expectations play an important role in treatment outcomes [1-4]. Although some studies suggested expectations of medical treatments are already too high and should be tempered by the clinician to cultivate realistic expectations for the patient [5-8], several meta-analyses have found that patients with more-positive pretreatment expectations achieve better outcomes [1-4]. Additionally, in patients treated for hand or wrist conditions, more-positive expectations have been reported to be associated with better outcomes [9-11]. In addition, positive expectations of the treatment outcomes are considered a key mechanism of placebo effects [12,13]. The placebo effect, or contextual nonspecific effect, is a psychobiological effect that is attributed to the overall therapeutic context [14,15]. This context can consist of patient-specific and clinician-specific factors, and the interaction of patient, clinician, treatment location, and treatment factors [16]. Clinical trials have shown considerable improvement in patients in placebo groups compared with an active or no treatment group [17,18]. Although positive expectations increase the contextual, nonspecific effects of a treatment, expectations may vary across patients and may depend on the type of treatment the patient is about to undergo. For example, previous studies showed that patients with hand or wrist disorders scheduled for surgery have higher expectations than similar patients scheduled for nonsurgical treatment [19,20].

Rationale

Using the contextual effects of a treatment may improve healthcare. Because the contextual nonspecific effect is believed to work through positive expectations of the outcome of a treatment, boosting expectations might be an important part of delivering high-quality care. However, little is known about factors independently associated with patient outcome expectations in clinical practice. Knowing the independent factors associated with outcome expectations may help clinicians to improve expectations. Improving expectations might, in turn, improve treatment outcomes. Moreover, it may inform future studies in the development of interventions that boost expectations.

Therefore, we asked: Which factors (such as, sociodemographics, patient-reported outcome measures [PROMs], illness perceptions, treatment, surgeon, and location) are independently associated with outcome expectations in patients with hand or wrist conditions?

Patients and Methods

Study Design

This was a cross-sectional study using a population-based sample of patients with hand or wrist conditions treated at our institution, and was reported following the STrengthening the Reporting of Observational studies in Epidemiology statement [21].

Setting

Data collection was part of usual care and occurred between July 2018 and December 2021 at Xpert Clinics. Xpert Clinics currently comprises 25 clinics for hand surgery and hand therapy in the Netherlands. Twenty-three surgeons are certified by the Federation of European Societies for Surgery of the Hand, and more than 150 hand therapists are employed at our treatment centers. Xpert Clinics offers insured care for hand and wrist conditions with no access restrictions because it is covered by public health insurance. At Xpert Clinics, outcomes are routinely evaluated [22]. After a diagnosis is registered during the first consultation, a measurement track is activated, and PROM forms are emailed to the patient. All data are digitally collected using GemsTracker electronic data capture tools (GemsTracker 2020, Erasmus MC and Equipe Zorgbedrijven), a secure internet-based application for distributing questionnaires and forms during clinical research and quality registrations. More details of the procedure at Xpert Clinics have been published [22].

Participants

Participants were eligible for inclusion if they were adults treated for a hand or wrist condition during the study period. We included patients from all measurement tracks, but excluded rare treatments with fewer than 20 patients for generalizability. Treatments can be divided into nonsurgical treatments (such as orthotics, exercise therapy, or injections), minor surgery (including trigger finger release or De Quervain release), and major surgery (such as trapeziectomy with or without ligament reconstruction tendon interposition for osteoarthritis of the thumb base, or corrective osteotomy for radius malunions). Additionally, we excluded patients who did not complete all relevant questionnaires. The number of patients treated during the study period determined the sample size.

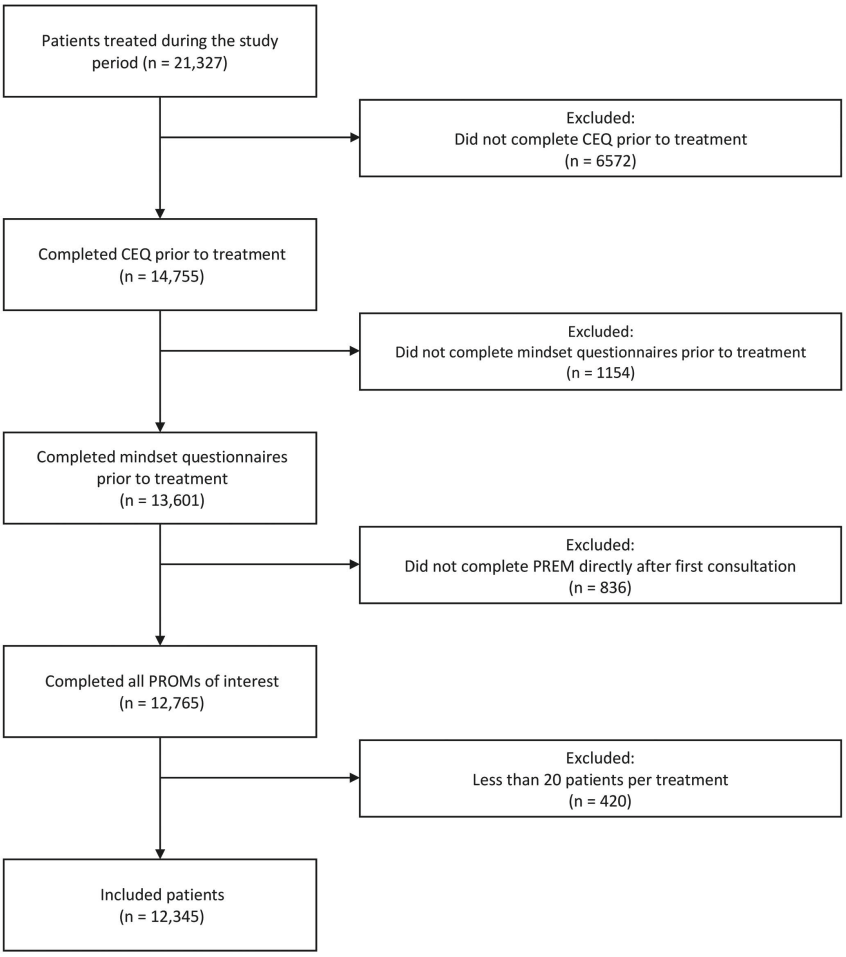


Fig. 1 This flowchart represents the patients who were included in this study. CEQ = Credibility and Expectancy Questionnaire.

We screened 21,327 patients with complete baseline sociodemographic data such as age and workload. Sixty percent (12,765 of 21,327) of patients completed all relevant PROMs. Finally, we e al sample (Fig. 1). To assess potential selection bias, we performed two nonresponder analyses. For this, we used the standardized mean difference as a measure of imbalance (standardized mean difference > 0.2 is considered to be imbalanced [23]). First, we compared the sociodemographic characteristics of patients who completed the Credibility/Expectancy Questionnaire (CEQ) (defined as responders) with patients who did not (defined as nonresponders). Second, we compared sociodemographic characteristics and the CEQ expectancy score of patients who additionally completed the other questionnaires of interest (responders) with patients who did not (nonresponders). In the first analysis, we found a small difference between responders and nonresponders (standardized mean difference = 0.43) (Supplemental Table 1; supplemental materials

are available with the online version of *CORR*®). In the second analysis, we found a small difference in treatment group (standardized mean difference = 0.28) and CEQ expectancy score (standardized mean difference = 0.21) (Supplemental Table 2; supplemental materials are available with the online version of *CORR*®). Those who participated were more likely to be in the surgical treatment group and to have higher expectations.

Table 1. Characteristics of the included patients (n = 12,345)

Characteristics	Total
Age in years	55 ± 15
Sex (female)	65 (7986)
Duration of symptoms in months	8 (4 -18)
Hand dominance	
Right	89 (10,960)
Left	8 (1013)
Both	3 (372)
Occupational intensity	
Not employed	37 (4553)
Light (working in an office)	28 (3506)
Moderate (working in a shop)	25 (3110)
Severe (working in construction)	10 (1176)
Second opinion	2 (301)
Recurrent disease	8 (1028)
Treatment group	
Nonsurgical treatment	29 (3544)
Minor surgery	49 (6022)
Major surgery	23 (2779)

Data presented as mean ± SD, median (IQR) or % (n).
Nonsurgical treatments includes e.g., orthotics, exercise therapy, injections; minor surgery includes minor surgical interventions e.g., trigger finger release, De Quervain release; major surgery includes more invasive interventions, e.g., trapeziectomy with or without ligament reconstruction tendon interposition for thumb base osteoarthritis, corrective osteotomy for radius malunions.

To assess the association between different degrees of surgical invasiveness, we distinguished nonsurgical treatment (such as hand therapy for thumb-base osteoarthritis), minor surgery (such as trigger finger release), and major surgery (such as Triangular Fibrocartilage Complex reinsertion). Twenty-nine percent (3544 of 12,345) of the final sample were scheduled for nonsurgical treatment, 49% (6022 of 12,345) for a minor surgical intervention, and 23% (2779 of 12,345) for a major surgical intervention (Table 1). The number of surgical patients in the present study does not reflect the actual distribution of surgical versus nonsurgical patients at Xpert Clinics, because the inclusion of patients in the present study depends on whether a measurement track is assigned. At the time of this study, no measurement tracks were started in our cohort in patients with, for example,

a “wait and see” policy or patients receiving steroid injections. Therefore, the proportion of surgical patients is overestimated in this study. Patients in the major surgery group had a longer duration of symptoms and were more often treated for the same disease previously. Patients in the minor surgery group had the most positive expectations (Supplemental Table 3; supplemental materials are available with the online version of *CORR*®). Furthermore, to assess potential differences between patients scheduled for nonsurgical treatment and patients scheduled for surgical treatment, we stratified patients into two treatment groups: nonsurgical and surgical. Seventy-one percent (8801 of 12,345) were scheduled for either minor or major surgery.

Variables and Measurements

The primary outcome in this study was patients’ outcome expectations of the treatment. We measured outcome expectations with the expectancy subscale of the CEQ [24]. This subscale consists of three items measuring the expected magnitude of improvement because of the prescribed treatment. Summed scores range from 3 to 27, where a higher score reflects a more positive treatment outcome expectation.

Independent Variables

We believed patients receiving the same treatment, counseled by the same surgeon, or treated at the same location might have more similar outcome expectations than other patients. To evaluate this, we used multilevel regression modeling with a random intercept and no fixed factors and intraclass correlation coefficients (ICC). Only for treatment, we found that patients were more similar in outcome expectations (Supplemental Digital Content 1; supplemental materials are available with the online version of *CORR*®). Therefore, we included the treatment level in all subsequent analyses.

Patient Characteristics

We divided patient characteristics into three subcategories: sociodemographics, PROMs for physical and mental health, and illness perception. Sociodemographic characteristics included age, sex (not gender, because we collect sex at the Dutch Citizen Service Administration, and we did not want to make unsupported assumptions), therapist-reported duration of symptoms (in months), hand dominance, therapist-reported occupational intensity (unemployed or light, moderate, or heavy physical labor), whether the patient visited the clinic for a second opinion, and whether the disease was recurrent (measured by the question: “Have been treated for the same disease before?”; the answer yes would be coded as recurrent. This means that a patient answering “yes” had the same or a different treatment for the same disease previously).

PROMs for physical and mental health included pain, hand function, health-related quality of life, psychologic distress, and pain catastrophizing at baseline. We used a VAS score (range 0 to 100) to measure the mean pain as experienced in the preceding week (higher scores indicate more pain) and hand function (higher scores indicate better function). The

VAS is a validated and widely used tool for measuring these constructs [25]. We measured health-related quality of life using the VAS of the EuroQol-5 Dimensions self-rated health questionnaire as an indication of the overall perceived health status (range 0 to 100; higher scores indicate better perceived health) [26,27]. Psychologic distress was measured with the Patient Health Questionnaire-4 (range 0 to 12; higher scores indicate more distress [28]), and pain catastrophizing was measured with the Pain Catastrophizing Scale (range 0 to 52; higher scores indicate a higher amount of catastrophizing [29]).

The last set of patient characteristics concerned illness perception as measured with the Brief Illness Perception Questionnaire [30,31]. The Brief Illness Perception Questionnaire measures patients' perception of their illness across eight domains (consequences, timeline, personal control, treatment control, identity, concern, coherence, and emotional response). Each domain is assessed with a single question (range 0 to 10; higher scores indicate more negative illness perceptions except for personal control, treatment control, and coherence, where the reverse is true) [30]. We excluded the domain of treatment control ("How much do you think your treatment can help your illness?") because of conceptual overlap with outcome expectations.

Treatment Characteristics

The treatment characteristics concerned the invasiveness and past effectiveness of the treatment. As an indicator of invasiveness, we coded a treatment as nonsurgical, minor surgery, or major surgery. In addition, as a proxy for the influence of the clinician's explanation of treatment effectiveness, for each treatment, we calculated the mean improvement in function achieved in patients treated previously, using VAS function scores (-100 = maximum deterioration in function; 100 = maximum improvement in function) administered at baseline and at 3 months. We did the same for pain (-100 = maximum deterioration in pain; 100 = maximum improvement in pain).

Finally, we used the Patient-Reported Experience Measure to measure the patient's experience with healthcare delivery, directly after the first consultation. This questionnaire is based on the Consumer Quality Index [32]. The Patient-Reported Experience Measure comprises 16 questions rated on a 4-point Likert scale, including questions about accessibility, reception in the clinic, and communication of the physician.

Ethical Approval

Ethical approval for this study was obtained from the medical ethics committee of the Erasmus MC Medical Centre, Rotterdam (MEC-2018-1088). Informed consent was obtained from patients before data collection started.

Statistical Methods

We used multilevel hierarchical regression analyses to test the relative association of specific patient and treatment characteristics with outcome expectations. In a hierarchical regression analysis, a set of variables is entered into a specific sequence to illustrate

each set's added amount of explained variance. This means that variables that add no or little to the explained variance remain in the model. In the first model, we entered all sociodemographic patient characteristics (such as sex, age, and occupational intensity). We added PROMs for physical and mental health (such as quality of life, pain, function, and psychologic distress) in the second model, illness perceptions in the third model, and treatment characteristics in the most-definitive model (the fourth model). An advantage of hierarchical regression is that because of shared variance, some variables might be pushed out of significance when entering the next step. Consequently, only variables that are truly associated with outcome expectations remain significant in the final model. For each model, the explained variance using multilevel partitioning was calculated.

Finally, we performed a stratified analysis to compare differences between factors associated with outcome expectations between patients scheduled for nonsurgical treatment and those scheduled for surgical treatment. Stratification is a useful strategy to identify interactions between subgroups such as treatment type.

A variance inflation factor greater than 3 was considered to indicate multicollinearity [33]. Based on the variance inflation factors (the highest-variance inflation factor in the multilevel hierarchical regression model equaled 2.05, in the stratified nonsurgical model, it equaled 2.12, and in the stratified surgical model, it equaled 2.03), we did not find any indication for multicollinearity in our models.

For all analyses, a p value < 0.05 was considered statistically significant. We used R statistical software version 4.1.1 for the analyses.

Results

Factors Independently Associated With Outcome Expectations

In our most-definitive model, we found an explained variance of 29%. When analyzing the separate steps of the different models, sociodemographics alone provided an explained variance of 1% in outcome expectations. PROMs for physical and mental health added 2% to the explained variance. Illness perceptions (9%) and treatment characteristics (18%) explained the largest amount of variance in outcome expectations.

Fig. 1 This flowchart represents the patients who were included in this study. CEQ = Credibility and Expectancy Questionnaire.

We used standardized betas to compare the magnitude of the effect of the different continuous and categorical variables. Higher outcome expectations were associated with the following sociodemographic variables (Fig. 2) (arranged from the largest to the smallest standardized beta coefficients): higher age (0.07; $p < 0.001$), occupational intensity (heavy: 0.06; $p = 0.02$; light: 0.06; $p = 0.002$; moderate: 0.06; $p = 0.008$), shorter duration of symptoms (0.03; $p < 0.001$); female sex (0.05; $p = 0.002$), and not having been treated for the same condition before (0.08; $p = 0.003$) (Table 2). Higher outcome expectations were associated with the following baseline PROMs for physical and mental health (largest to smallest standardized beta coefficients): a higher EQ-5D self-rated health score (0.07; $p < 0.001$), better hand function (0.05; $p < 0.001$), and more pain catastrophizing (0.02; $p = 0.048$). Six of seven illness perception items were associated with greater outcome expectations (from largest to smallest): shorter illness duration expected by the patient (-0.23; $p < 0.001$), better understanding of the condition by the patient (0.12; $p < 0.001$), the more the illness affects the patient's life (0.09; $p < 0.001$), less

concern about the illness the patient experiences (-0.08; $p < 0.001$), a larger number of symptoms the patient views as being part of their illness (0.05; $p < 0.001$), and the less the illness affects the patient emotionally (-0.04; $p < 0.001$). The largest standardized beta coefficients were for treatment characteristics: major surgical treatment (0.61; $p < 0.001$) and minor surgical treatment (0.56; $p < 0.001$). This means that being at the start of a major surgical treatment increases the outcome expectations by 2.75 points (95% confidence interval 2.21 to 3.29; $p < 0.001$) compared with being at the start of a nonsurgical treatment (Supplemental Table 4; supplemental materials are available with the online version of *CORR*®). The mean functional improvement of the treatment was also associated with outcome expectations (0.17; $p < 0.001$) (Table 2).

Analyzing differences in variables between the different steps of the model, we found only one difference (Supplemental Table 5; supplemental materials are available with the online version of *CORR*®). In Model 1, visiting the clinic for a second opinion was associated with lower expectations, but after adding PROMs for physical and mental health, there was no association. This implies that one (or more) of the PROMs, such as pain catastrophizing, have a shared variance with a second opinion and pushes the variable second opinion out of significance.

Table 2. Most-definitive model (standardized beta coefficients) after the hierarchical linear regression analyses (n = 12,345) using sociodemographics, PROMs for physical and mental health, illness perception, and treatment characteristics explaining outcome expectations

	Expectations for all treatments		Expectations for nonsurgical treatment		Expectations for surgical treatment	
Variables	**Standardized coefficients (95% CI)**	**p value**	**Standardized coefficients (95% CI)**	**p value**	**Standardized coefficients (95% CI)**	**p value**
Sociodemographics						
Age in years	0.07 (0.05 to 0.09)	< 0.001	0.07 (0.04 to 0.11)	< 0.001	0.08 (0.06 to 0.11)	< 0.001
Sex (female)	0.05 (0.02 to 0.09)	0.002	0.09 (0.03 to 0.16)	0.01	0.05 (0.0 to 0.09)	0.04
Light occupational intensity (reference: not employed)	0.06 (0.02 to 0.10)	0.002	0.08 (0.00 to 0.16)	0.04	0.06 (0.00 to 0.11)	0.04
Moderate occupational intensity (reference not employed)	0.06 (0.02 to 0.10)	0.01	0.09 (0.01 to 0.17)	0.03	0.04 (-0.02 to 0.09)	0.18
Heavy occupational intensity (reference not employed)	0.06 (0.01 to 0.12)	0.02	0.05 (-0.07 to 0.16)	0.43	0.09 (0.01 to 0.16)	0.02
Second opinion: No	0.06 (-0.04 to 0.16)	0.21	0.06 (-0.14 to 0.26)	0.57	0.04 (-0.08 to 0.17)	0.49
Duration of symptoms in months	-0.03 (-0.05 to -0.02)	< 0.001	-0.04 (-0.07 to -0.01)	0.004	-0.03 (-0.05 to -0.01)	0.01
Right hand dominance (reference: left)	-0.01 (-0.06 to 0.04)	0.7	-0.04 (-0.15 to 0.07)	0.49	-0.01 (-0.08 to 0.06)	0.78
Both hand dominance (reference: left)	-0.03 (-0.13 to 0.07)	0.56	-0.07 (-0.26 to 0.13)	0.5	-0.02 (-0.15 to 0.11)	0.77
Recurrent: yes	-0.08 (-0.14 to -0.03)	0.003	-0.12 (-0.28 to 0.05)	0.17	-0.11 (-0.18 to -0.05)	0.001
PROMs for physical and mental health						
VAS function	0.05 (0.03 to 0.06)	< 0.001	0.03 (0.00 to 0.06)	0.05	0.06 (0.04 to 0.09)	< 0.001
VAS pain	0.01 (-0.01 to 0.03)	0.21	0.01 (-0.02 to 0.05)	0.41	0.01 (-0.01 to 0.04)	0.29

Table 2. *(continued)*

	Expectations for all treatments		Expectations for nonsurgical treatment		Expectations for surgical treatment	
Variables	**Standardized coefficients (95% CI)**	**p value**	**Standardized coefficients (95% CI)**	**p value**	**Standardized coefficients (95% CI)**	**p value**
EQ-5D self-rated health	0.07 (0.05 to 0.09)	< 0.001	0.09 (0.05 to 0.13)	< 0.001	0.07 (0.04 to 0.10)	< 0.001
Pain Catastrophizing Score	0.02 (0.00 to 0.04)	0.048	0.01 (-0.02 to 0.05)	0.46	0.03 (0.00 to 0.05)	0.03
PHQ Depression Score	-0.01 (-0.04 to 0.01)	0.18	-0.04 (-0.09 to -0.00)	0.04	0.00 (-0.02 to 0.03)	0.74
PHQ Anxiety Score	0.01 (-0.01 to 0.03)	0.3	0.08 (0.04 to 0.12)	< 0.001	-0.03 (-0.05 to 0.00)	0.06
Illness perception						
B-IPQ Consequences	0.09 (0.07 to 0.11)	< 0.001	0.11 (0.07 to 0.15)	< 0.001	0.10 (0.07 to 0.13)	< 0.001
B-IPQ Timeline	-0.23 (-0.24 to -0.21)	< 0.001	-0.36 (-0.39 to -0.32)	< 0.001	-0.21 (-0.24 to -0.19)	< 0.001
B-IPQ Personal control	0.01 (-0.01 to 0.02)	0.45	0.13 (0.10 to 0.16)	< 0.001	-0.05 (-0.07 to -0.03)	< 0.001
B-IPQ Identity	0.05 (0.03 to 0.06)	< 0.001	0.03 (-0.01 to 0.06)	0.15	0.06 (0.04 to 0.09)	<0.001
B-IPQ Concern	-0.08 (-0.10 to 0.06)	< 0.001	-0.06 (-0.10 to -0.02)	0.002	-0.10 (-0.13 to -0.07)	< 0.001
B-IPQ Coherence	0.12 (0.10 to 0.13)	< 0.001	0.12 (0.09 to 0.15)	< 0.001	0.13 (0.11 to 0.15)	< 0.001
B-IPQ Emotional response	-0.04 (-0.06 to -0.02)	< 0.001	-0.02 (-0.06 to 0.02)	0.23	-0.07 (-0.10 to -0.04)	< 0.001
Treatment characteristics						
Type treatment (minor surgery)	0.56 (0.44 to 0.68)	< 0.001	NA	NA	NA	NA
Type treatment (major surgery)	0.61 (0.49 to 0.73)	< 0.001	NA	NA	0.01 (-0.11 to 0.12)	0.92
Mean improvement pain	-0.05 (-0.11 to 0.02)	0.167	-0.02 (-0.08 to 0.05)	0.62	-0.02 (-0.09 to 0.05)	0.59
Mean improvement function	0.17 (0.11 to 0.24)	< 0.001	0.15 (0.09 to 0.22)	< 0.001	0.11 (0.04 to 0.18)	0.002
Random effects						
σ^2	0.67		0.75		0.85	

Table 2. *(continued)*

	Expectations for all treatments		Expectations for nonsurgical treatment		Expectations for surgical treatment	
Variables	**Standardized coefficients (95% CI)**	**p value**	**Standardized coefficients (95% CI)**	**p value**	**Standardized coefficients (95% CI)**	**p value**
$\tau_{00\ treatment}$	0.02		0.00		0.02	
ICC	0.03		0.01		0.02	
$N_{treatment}$	56		17		39	
Observations	12,345		3544		8801	
Marginal or conditional r^2	0.293/0.314		0.252/0.256		0.137/0.154	

Standardized beta coefficients, 95% CIs, and p values are displayed, along with the random effects and explained variance expressed in the marginal r^2 for the most-definitive model with all treatments, nonsurgical treatments, and surgical treatments. Standardized β coefficients, converted to the same scale, are reported to allow easier between-variable comparisons and determine the relative association of each explanatory variable. EQ-5D = EuroQol-5 Dimensions; PROM = patient-reported outcome measures; B-IPQ = Brief Illness Perception Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire.

Differences Between Patients Scheduled for Nonsurgical Treatment and Those Scheduled for Surgical Treatment

In the most-definitive model, including sociodemographics, PROMs for physical and mental health, illness perception, and treatment characteristics, we found an explained variance of 25% for outcome expectations of patients scheduled for nonsurgical treatment. Sociodemographics explained 2%, PROMs for physical and mental health explained 2%, illness perception explained 16%, and treatment characteristics explained 5%. For the outcome expectations of patients scheduled for surgical treatment, the most-definitive model explained 14% of the variance. Sociodemographics explained 2%, PROMs explained 2%, illness perception explained 8%, and treatment characteristics explained 2%.

When comparing the factors associated with outcome expectations between patients scheduled for nonsurgical treatment and those scheduled for surgical treatment, we found greater personal control was associated with more-positive expectations in nonsurgical patients (0.13; $p < 0.001$), whereas higher personal control was associated with more-negative expectations in surgical patients (-0.05; $p < 0.001$) (Fig. 3). Psychologic distress was associated with expectations only in nonsurgical patients (depression: -0.04; $p = 0.04$; anxiety: 0.08; $p < 0.001$). Pain catastrophizing (0.03; $p = 0.03$), whether the patient has been treated for the same disease before (-0.11; $p = 0.001$), and a larger number of symptoms the patient views as being part of their Illness (0.08; $p < 0.001$) were associated with expectations only in surgical patients (Table 2).

Discussion

Multiple studies have shown that more-positive outcome expectations are associated with better treatment outcomes [1-4,9-11], and there is an interest in positively modifying outcome expectations to improve treatment outcomes. However, little was known about factors independently associated with outcome expectations. Studying factors associated with outcome expectations may provide relevant information for clinicians and researchers aiming to improve outcome expectations. Improving expectations might, in turn, improve treatment outcomes. We found patients' outcome expectations for a hand or wrist condition were higher when patients had more-positive perceptions of their illness. Furthermore, patients scheduled for surgical treatment had higher outcome expectations than patients scheduled for nonsurgical treatment, even after adjusting for differences in clinical profile and mindset between patents. Our findings can be used directly in daily clinic by improving expectations and illness perceptions, especially for nonsurgical patients, or in studies that develop interventions to improve expectations.

Fig. 3 Standardized regression coefficients of the stratified hierarchical multilevel regression models explain outcome expectations for (A) nonsurgical treatment and (B) surgical treatment. Only significant variables are shown. EQ-5D = EuroQol-5 Dimensions; B-IPQ = Brief Illness Perception Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire.

Limitations

Our study has several limitations. First, because this was an observational study, no causal conclusions can be drawn. Although we theorized the variables in our model drive outcome expectations, the reverse could be just as true for several variables (outcome expectations may be causing illness perceptions), or instead, the relationship may be bidirectional. Experimental studies are necessary to test whether outcome expectations might be strengthened by influencing illness perceptions. Second, we found small differences between patients who responded to the survey (responders) and those who did not (nonresponders). Nonresponders were more often scheduled for nonsurgical treatment and had lower expectations. This is in line with other studies that showed nonsurgical patients are more likely to be lost to follow-up than surgical patients [9,20,34]. Furthermore, our study and others showed that patients scheduled for nonsurgical treatment have lower expectations [7,20,35], so we may assume the difference in expectations between responders and nonresponders is caused by the difference in treatment type we found in the nonresponder analysis. Still, we may have overestimated the expectations of patients undergoing nonsurgical treatment in our study. Third, our study examined pretreatment expectations, but several studies suggested outcome expectations may change during treatment and this change may influence treatment outcomes [35,36]. Nevertheless, a robust association between pretreatment outcome expectations and treatment outcomes has been found in several medical areas [1, 5, 6, 27], indicating the importance of addressing pretreatment expectations. Future research could investigate whether the extent to which outcome expectations change during treatment depends on the type of treatment and how this change affects outcome.

Association of Location, Surgeon, and Treatment Variables With Outcome Expectations

Nineteen percent of the variance in outcome expectations was attributable to differences between treatments rather than differences within treatments. Considering the surgeon and location level, we found the variance in outcome expectations was because of differences in surgeon or location, and almost none was because of differences between surgeons or locations. Theoretically, a surgeon adjusts his or her behavior to the patient, treatment, or other factors, such as workload. This might explain why we mainly saw within-surgeon differences.

Patient and Treatment Factors Independently Associated With Outcome Expectations

Our study showed illness perception is an important factor strongly associated with outcome expectations. The more positively patients perceived their illnesses, the more positive their expectations were of the treatment outcome. Perceived chronicity of the disease and the perceived understanding of the disease displayed the strongest independent association. Given studies usually investigate variables associated with outcome expectations of a single treatment, previous researchers may have missed an

important overarching factor driving expectations: the type of treatment a patient is about to undergo. In our study, approximately 18% of the total variance across patients was explained by the treatment invasiveness (nonsurgical, minor, or major surgical treatments) and the past effectiveness of the treatment. These results might indicate that patients believe treatment invasiveness is positively associated with better outcomes, resulting in higher pretreatment expectations by patients scheduled for surgical treatment. This finding is in line with those of other studies [19,20,34]. Our study indicates that expectation management should be tailored to the specific treatment (surgical or nonsurgical) and to the specific patient (including their perception of illness). For example, an intervention aimed to increase the understanding of a specific illness and accompanying treatment (such as offering an illness-specific or patient-specific elearning module with psychoeducation to provide information and support so a patient will better understand their illness and treatment) might effectively correct false (negative) beliefs regarding treatment invasiveness in nonsurgical patients and thus improve their pretreatment expectations.

We found an association with the treatment effectiveness based on the mean improvement in function in historical patients, but not for the mean improvement in pain. Hypothetically, in their explanation of treatment effectiveness, clinicians might avoid strong statements about pain, because the amount of improvement in pain differs greatly between patients and between treatments. However, statements on hand function, including a statement such as: “you will be able to return to work within 12 weeks,” might be safer because this outcome may be more predictable. Additionally, we did not find an association between the amount of pain at baseline, whereas for function, we found patients with better pretreatment function had higher expectations. This suggests pain might be less important for outcome expectations than pretreatment level of function is.

Differences Between Patients Scheduled for Nonsurgical Treatment and Those Scheduled for Surgical Treatment

The degree of control patients feel they have over their illness was the only illness perception domain not associated with outcome expectations in our hierarchical regression model. However, our stratified analysis shows that the more personal control a nonsurgical patient experienced, the more positive the outcome expectations were, whereas the reverse was true for surgical patients. Because of this opposite effect, they may likely have cancelled each other out in the overall regression analysis. This opposite effect might guide intervention for improving outcome expectations. Patients with an internal locus of control perceive themselves as having a great deal of personal control over their outcomes, whereas patients with an external locus of control believe their outcomes result from external influences. Considering the locus of control, improving outcome expectations in nonsurgical patients should entail an increase in personal control (such as a greater understanding of illness and self-efficacy). In contrast, the outcome

expectations of surgical patients might be improved by discussing important external influences (including physician experience and the likelihood of success with treatment).

Conclusion

So far, there is some promising evidence for expectancy-focused interventions to improve treatment outcomes [37]. Expectation management appears to be an important element of delivering high-quality healthcare. Our findings suggest expectation management should be tailored to the specific treatment (such as surgical versus nonsurgical) and the specific patient (including their perception of their illness). It may be more beneficial to test and implement expectation management strategies such as physical therapy for nonsurgical treatments than for surgical treatments, given our findings indicate a greater need to do so. An additional advantage of such a strategy is that successful interventions may be able to prevent converting to surgical interventions, which is a goal of the stepped-care principles of standard care. Future studies might investigate the association between pretreatment expectations and outcomes by performing an experimental study, such as a randomized controlled trial, in which boosting expectations is compared with usual care (with no special attention to expectations) in nonsurgical and surgical groups.

Acknowledgments

We thank all patients who participated and allowed their data to be anonymously used for the present study.

Group Authors

Members of the Hand-Wrist Study Group include: Richard Arjen Michiel Blomme, Dirk-Johannes Jacobus Cornelis van der Avoort, Alexander Kroeze, Jeronimus Maria Smit, Jan Debeij, Erik Taco Walbeehm, Gijs Marijn van Couwelaar, Johannes Pieter de Schipper, Johannes Frederikes Maria Temming, Jeroen Hein van Uchelen, Herman Luitzen de Boer, Kennard Harmsen, Oliver Theodor Zöphel, Thybout Matthias Moojen, Xander Smit, Gert-Jan Halbesma, Rob van Huis, Pierre Yves Alain Adriaan Pennehouat, Karin Schoneveld, Yara Eline van Kooij, Joris Jan Veltkamp, Alexandra Fink, Joris Sebastiaan Teunissen, Jaimy Emerentiana Koopman, Nina Louisa Loos, Jak Dekker, Ward Rogier Bijlsma, and Marloes Hendrina Paulina ter Stege

References

1. Auer CJ, Glombiewski JA, Doering BK, et al. Patients' expectations predict surgery outcomes: a meta-analysis. *Int J Behav Med.* 2016;23:49-62.
2. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res.* 2006;60:631-637.
3. Charlesworth JEG, Petkovic G, Kelley JM, et al. Effects of placebos without deception compared with no treatment: A systematic review and meta-analysis. *J Evid Based Med.* 2017;10:97-107.
4. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988
5. Constantino MJ, Glass CR, Arnkoff DB, Ametrano RM, Smith JZ. Expectations. *Psychotherapy Relationships That Work: Evidence-Based Responsiveness*: Oxford University Press; 2011.
6. Crow R, Gage H, Hampson S, et al. The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technol Assess.* 1999;3:1-96.
7. de Raaij EJ, Schröder C, Maissan FJ, Pool JJ, Wittink H. Cross-cultural adaptation and measurement properties of the Brief Illness Perception Questionnaire-Dutch Language Version. *Man Ther.* 2012;17:330-335.
8. De Ridder WA, Wouters RM, Hoogendam L, et al. Which factors are associated with satisfaction with treatment results in patients with hand and wrist conditions? A large cohort analysis. *Clin Orthop Relat Res.* 2022.
9. Delnoij DM, Rademakers JJ, Groenewegen PP. The Dutch consumer quality index: an example of stakeholder involvement in indicator development. *BMC Health Serv Res.* 2010;10:88.
10. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry.* 2000;31:73-86.
11. Duivenvoorden T, Vissers MM, Verhaar JA, et al. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study. *Osteoarthritis Cartilage.* 2013;21:1834-1840.
12. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Biological, clinical, and ethical advances of placebo effects. *Lancet.* 2010;375:686-695.
13. Frouzakis R, Herren DB, Marks M. Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis. *J Hand Surg Am.* 2015;40:483-490.
14. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken).* 2011;63 Suppl 11:S240-252.
15. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20:1727-1736.
16. Hoogendam L, Koopman JE, van Kooij YE, et al. What are the minimally important changes of four commonly used patient-reported outcome measures for 36 hand and wrist condition-treatment combinations? *Clin Orthop Relat Res.* 2022;480:1152-1166.

17. Hoogendam L, van der Oest MJW, Wouters RM, et al. Patients with higher treatment outcome expectations are more satisfied with the results of nonoperative treatment for thumb base osteoarthritis: a cohort study. *Arch Phys Med Rehabil.* 2021;102:1533-1540.

18. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *Pharmacoeconomics.* 2018;36:675-697.

19. Kadzielski J, Malhotra LR, Zurakowski D, et al. Evaluation of preoperative expectations and patient satisfaction after carpal tunnel release. *J Hand Surg Am.* 2008;33:1783-1788.

20. Kong J, Benedetti F. Placebo and nocebo effects: an introduction to psychological and biological mechanisms. *Handb Exp Pharmacol.* 2014;225:3-15.

21. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* 2009;50:613-621.

22. Kutner MHN, C.J., Neter, J., Li, W. *Applied linear statistical models.* Boston: McGraw-Hill; 2005

23. Mandl LA, Burke FD, Shaw Wilgis EF, et al. Could preoperative preferences and expectations influence surgical decision making? Rheumatoid arthritis patients contemplating metacarpophalangeal joint arthroplasty. *Plast Reconstr Surg.* 2008;121:175-180.

24. Marks M, Audigé L, Reissner L, et al. Determinants of patient satisfaction after surgery or corticosteroid injection for trapeziometacarpal osteoarthritis: Results of a prospective cohort study. *Arch Orthop Trauma Surg.* 2015;135:141-147.

25. Miller FG, Kaptchuk TJ. The power of context: reconceptualizing the placebo effect. *J R Soc Med.* 2008;101:222-225.

26. Myers H, Keefe F, George SZ, et al. The influence of a cognitive behavioural approach on changing patient expectations for conservative care in shoulder pain treatment: a protocol for a pragmatic randomized controlled trial. *BMC Musculoskelet Disord.* 2021;22:727.

27. Peerdeman KJ, van Laarhoven AIM, Keij SM, et al. Relieving patients' pain with expectation interventions: a meta-analysis. *Pain.* 2016;157:1179-1191.

28. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol.* 2008;59:565-590.

29. Selles RW, Wouters RM, Poelstra R, et al. Routine health outcome measurement: development, design, and implementation of the Hand and Wrist Cohort. *Plast Reconstr Surg.* 2020;146:343-354.

30. Stewart-Williams S, Podd J. The placebo effect: dissolving the expectancy versus conditioning debate. *Psychol Bull.* 2004;130:324-340.

31. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment.* 1995;7:524-532.

32. Tolk JJ, Janssen RPA, Haanstra TM, et al. The influence of expectation modification in knee arthroplasty on satisfaction of patients: a randomized controlled trial. *Bone Joint J.* 2021;103-B:619-626.

33. van der Oest MJW, Hoogendam L, Wouters RM, et al. Associations between positive treatment outcome expectations, illness understanding, and outcomes: a cohort study on non-operative treatment of first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021:1-8.

34. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol.* 2008;61:344-349.

35. Waljee JF, Chung KC. Commentary regarding "Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis". *J Hand Surg Am.* 2015;40:491-492.

36. Wouters RM, Vranceanu AM, Slijper HP, et al. Patients with thumb-base osteoarthritis scheduled for surgery have more symptoms, worse psychological profile, and higher expectations than nonsurgical counterparts: a large cohort analysis. *Clin Orthop Relat Res.* 2019;477:2735-2746.

37. Zou K, Wong J, Abdullah N, et al. Examination of overall treatment effect and the proportion attributable to contextual effect in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis.* 2016;75:1964-1970.

Supplementary Table 1. First nonresponder analysis.

Characteristics	Responders (n = 12,765)	Nonresponders (n = 8562)	Standardized Mean Difference
Age in years (mean)	55 ± 15	53 ± 17	0.11
Sex (female)	65 (8231)	64 (5448)	0.02
Duration of symptoms in months (median)	8 (4-18)	6 (3-14)	0.02
Hand dominance			0.04
Right	89 (11,340)	90 (7671)	
Left	8 (1047)	8 (687)	
Both	3 (378)	2 (204)	
Occupational intensity			0.04
Not employed	37 (4697)	37 (3176)	
Light (working in an office)	28 (3630)	27 (2343)	
Moderate (working in a shop)	25 (3217)	25 (2129)	
Severe (working in construction)	10 (1221)	11 (914)	
Second opinion			0.01
Yes	3 (344)	2.5 (215)	
Recurrent disease			0.06
Yes	8 (1065)	7 (583)	
Treatment group			0.43
Nonsurgical treatment	28 (3625)	48 (4113)	
Minor surgery	48 (6091)	38 (3262)	
Major surgery	24 (3049)	14 (1187)	

Data presented as mean ± SD, % (n), or median (IQR). Patient characteristics for patients who completed all questionnaires of interest (responders), before exclusion of treatments with less than 20 patients, compared with patients who did not complete all questionnaires of interest (nonresponders). The Standardized Mean Difference is used as an indication of imbalance (SMD > 0.2 is considered to be imbalanced).

Supplementary Table 2. Second nonresponder analysis.

Characteristics	Responders (n = 12,765)	Nonresponders (n = 1990)	Standardized Mean Difference
Age in years (mean)	55 ± 15	55 ± 16	0.02
Sex **(**female)	65 (8231)	67 (1333)	0.05
Duration of symptoms in months (median)	8 (4-18)	6 (3-18)	0.02
Hand dominance			0.04
Right	89 (11,340)	89 (1763)	
Left	8 (1047)	9 (178)	
Both	3 (378)	3 (49)	
Occupational intensity			0.04
Not employed	37 (4697)	38 (764)	
Light (working in an office)	28 (3630)	27 (535)	

Supplementary Table 2. Second nonresponder analysis. *(continued)*

Characteristics	Responders (n = 12,765)	Nonresponders (n = 1990)	Standardized Mean Difference
Moderate (working in a shop)	25 (3217)	25 (497)	
Severe (working in construction)	10 (1221)	10 (194)	
Second opinion (yes)	3 (344)	3 (65)	0.03
Recurrent disease (yes)	8 (1065)	7 (148)	0.03
Treatment group			0.28
Nonsurgical treatment	28 (3625)	41 (822)	
Minor surgery	48 (6091)	41 (811)	
Major surgery	24 (3049)	18 (357)	
CEQ Expectancy score	22 ± 5	21 ± 5	0.21

Data presented as mean ± SD, % (n), or median (IQR). Patient characteristics for patients who completed all questionnaires of interest (responders), before exclusion of treatments with less than 20 patients, compared with patients who did complete the Credibility and Expectancy Questionnaire, but did not complete the other questionnaires of interest (nonresponders). The Standardized Mean Difference is used as an indication of imbalance.

Supplementary table 3. Characteristics of the included patients per treatment group

Characteristics	Nonsurgical (n=3544)	Minor surgery (n=6022)	Major surgery (n=2779)
Age in years, mean ± SD	53 ± 16	56 ± 15	56 ± 15
Female sex, % (n)	72 (2535)	66 (3995)	52 (1456)
Duration of symptoms in months, median (IQR)	6 (3-12)	7 (4-12)	12 (6-28)
Hand dominance, % (n)			
Right	90 (3178)	89 (5373)	87 (2409)
Left	7 (258)	8 (489)	10 (266)
Both	3 (308)	3 (160)	4 (104)
Occupational intensity, % (n)			
Not employed	34 (1191)	38 (2283)	39 (1079)
Light (working in an office)	30 (1064)	27 (1640)	29 (802)
Moderate (working in a shop)	28 (981)	25 (1519)	22 (610)
Severe (working in a shop)	9 (308)	10 (580)	10 (288)
Second opinion, % (n)	2 (74)	2 (93)	5 (134)
Recurrent disease, % (n)	3 (110)	9 (534)	14 (384)
CEQ expectations score, mean ± SD	19 ± 5	23 ± 4	22 ± 3

Nonsurgical treatments includes e.g., orthotics, exercise therapy, injections; minor surgery includes minor surgical interventions e.g., trigger finger release, De Quervain release; major surgery includes more invasive interventstions, e.g., trapeziectomy with or without ligament reconstruction tendon interposition for thumb base osteoarthritis, corrective osteotomy for radius malunions. CEQ = Credibility and Expectancy Questionnaire.

Supplementary Table 4. Most-definitive model (nonstandardized beta coefficients) after the hierarchical linear regression analyses (n = 12,345) using sociodemographics, Patient Reported Outcome Measures on physical and mental health, illness perception, and treatment characteristics explaining outcome expectations.

	Expectations total		Expectations nonsurgical		Expectations surgical	
Variables	**Nonstandardized coefficient (95% CI)**	**p-value**	**Nonstandardized coefficient (95% CI)**	**p-value**	**Nonstandardized coefficient (95% CI)**	**p-value**
Sociodemographics						
Age in years	0.02 (0.01 to 0.03)	< 0.001	0.02 (0.01 to 0.04)	< 0.001	0.02 (0.01 to 0.03)	< 0.001
Sex (male)	-0.24 (-0.39 to -0.09)	0.002	-0.49 (-0.83 to -0.14)	0.006	-0.16 (-0.32 to -0.01)	0.04
Light occupational intensity (reference: not employed)	0.29 (0.11 to 0.46)	0.002	0.42 (0.02 to 0.82)	0.04	0.20 (0.01 to 0.38)	0.04
Moderate occupational intensity (reference: not employed)	0.26 (0.07 to 0.44)	0.008	0.47 (0.05 to 0.88)	0.03	0.13 (-0.06 to 0.33)	0.18
Heavy occupational intensity (reference: not employed)	0.29 (0.04 to 0.55)	0.02	0.24 (-0.35 to 0.83)	0.43	0.30 (0.04 to 0.57)	0.02
Second Opinion: No	0.28 (-0.16 to 0.71)	0.21	0.30 (-0.74 to 1.34)	0.57	0.16 (-0.29 to 0.60)	0.49
Duration of symptoms in months	-0.00 (-0.01 to -0.00)	< 0.001	-0.01 (-0.01 to -0.00)	0.004	-0.00 (-0.00 to -0.00)	0.01
Right hand dominance (reference: left)	-0.05 (-0.28 to 0.19)	0.7	-0.20 (-0.77 to 0.37)	0.49	-0.03 (-0.28 to 0.21)	0.78
Both hand dominance (reference: left)	-0.13 (-0.57 to 0.31)	0.56	-0.35 (-1.36 to 0.67)	0.5	-0.07 (-0.53 to 0.39)	0.77
Recurrent: yes	-0.38 (-0.62 to -0.13)	0.003	-0.61 (-1.47 to 0.26)	0.17	-0.40 (-0.62 to -0.17)	0.001
PROMs on physical and mental health						
VAS Function	0.01 (0.01 to 0.01)	< 0.001	0.01 (0.00 to 0.01)	0.047	0.01 (0.01 to 0.01)	< 0.001
VAS pain	0.00 (-0.00 to 0.01)	0.21	0.00 (-0.00 to 0.01)	0.41	0.00 (-0.00 to 0.01)	0.29
EQ-5D self-rated health	1.63 (1.18 to 2.08)	< 0.001	2.49 (1.46 to 3.53)	< 0.001	1.22 (0.76 to 1.69)	< 0.001
Pain Catastrophizing Score	0.01 (0.00 to 0.02)	0.048	0.01 (-0.01 to 0.03)	0.46	0.01 (0.00 to 0.02)	0.03

Supplementary Table 4. *(continued)*

	Expectations total		Expectations nonsurgical		Expectations surgical	
Variables	**Nonstandardized coefficient (95% CI)**	**p-value**	**Nonstandardized coefficient (95% CI)**	**p-value**	**Nonstandardized coefficient (95% CI)**	**p-value**
PHQ Depression Score	-0.06 (-0.14 to 0.03)	0.18	-0.19 (-0.37 to -0.01)	0.04	0.01 (-0.07 to 0.10)	0.74
PHQ Anxiety Score	0.04 (-0.03 to 0.11)	0.3	0.31 (0.15 to 0.47)	< 0.001	-0.07 (-0.15 to 0.00)	0.06
Illness Perception						
B-IPQ Consequences	0.16 (0.12 to 0.19)	< 0.001	0.24 (0.16 to 0.32)	< 0.001	0.13 (0.09 to 0.17)	< 0.001
B-IPQ Timeline	-0.37 (-0.40 to -0.34)	< 0.001	-0.72 (-0.78 to -0.65)	< 0.001	-0.27 (-0.30 to -0.24)	< 0.001
B-IPQ Personal control	0.01 (-0.02 to 0.04)	0.45	0.31 (0.23 to 0.38)	< 0.001	-0.06 (-0.09 to -0.04)	< 0.001
B-IPQ Identity	0.08 (0.05 to 0.11)	< 0.001	0.06 (-0.02 to 0.13)	0.15	0.08 (0.05 to 0.12)	< 0.001
B-IPQ Concern	-0.13 (-0.16 to -0.10)	< 0.001	-0.12 (-0.20 to -0.04)	0.002	-0.12 (-0.15 to -0.08)	< 0.001
B-IPQ Coherence	0.24 (0.21 to 0.27)	< 0.001	0.27 (0.20 – 0.34)	< 0.001	0.22 (0.19 to 0.26)	< 0.001
B-IPQ Emotional response	-0.06 (-0.09 to -0.03)	< 0.001	-0.04 (-0.12 to 0.03)	0.23	-0.08 (-0.11 to -0.05)	< 0.001
Treatment characteristics						
Type treatment (minor surgery)	2.53 (1.97 to 3.10)	< 0.001	NA	NA	NA	NA
Type treatment (major surgery)	2.75 (2.21 to 3.29)	< 0.001	NA	NA	0.02 (-0.37 to 0.41)	0.92
Mean improvement pain	-0.02 (-0.05 to 0.01)	0.17	-0.02 (-0.08 to 0.05)	0.62	-0.01 (-0.03 to 0.02)	0.59
Mean improvement function	0.11 (0.07 to 0.15)	< 0.001	0.16 (0.09 to 0.22)	< 0.001	0.06 (0.02 to 0.10)	0.002
Random effects						
σ^2	13.67		20.06		10.57	
$\tau_{00\ treatment}$	0.42		0.11		0.22	
ICC	0.03		0.01		0.02	
$N_{\ treatment}$	56		17		39	

Supplementary Table 4. *(continued)*

	Expectations total		Expectations nonsurgical		Expectations surgical	
Variables	**Nonstandardized coefficient (95% CI)**	**p-value**	**Nonstandardized coefficient (95% CI)**	**p-value**	**Nonstandardized coefficient (95% CI)**	**p-value**
Observations	12345		3544		8801	
Marginal R^2/conditional R^2	0.293/0.314		0.252/0.256		0.137/0.154	

Nonstandardized beta coefficients, 95% CIs, and p values are displayed, along with the random effects and explained variance expressed in the Marginal R^2 for the most-definitive model with (1) all treatments, (2) nonsurgical treatments, (3) surgical treatments. The nonstandardized estimates in our most definitive model indicate that with every unit increase in either a continuous, dichotomous, or categorical independent variable, the outcome expectations increase or decrease by the value of the nonstandardized estimate. EQ-5D = EuroQol-5 Dimensions; B-IPQ = Brief Illness Perception Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire.

Supplemental Table 5. Standardized and nonstandardized beta coefficients for the hierarchical multilevel models explaining expectations before treatment.

	Model 1		Model 2	
Explanatory variables	**B (95% CI)**	**β (95% CI)**	**B (95% CI)**	**β (95% CI)**
Sociodemographics				
Age in years	0.03 (0.02 to 0.04[c]	0.10 (0.08 to 0.12)[c]	0.02 (0.02 to 0.03)[c]	0.08 (0.06 to 0.10)[c]
Male sex	-0.10 (-0.26 to 0.06)	-0.02 (-0.06 to 0.01)	-0.21 (-0.36 to -0.05)[a]	-0.05 (-0.08 to -0.01)[a]
Workload (unemployed)				
Light	0.45 (0.27 to 0.64)[c]	0.10 (0.06 to 0.14)[c]	0.26 (0.08 to 0.45)[b]	0.06 (0.02 to 0.10)[b]
Moderate	0.45 (0.25 to 0.64)[c]	0.10 (0.06 to 0.14)[c]	0.33 (0.14 to 0.53)[c]	0.07 (0.03 to 0.12)[c]
Heavy	0.52 (0.26 to 0.79)[c]	0.12 (0.06 to 0.18)[c]	0.44 (0.18 to 0.71)[c]	0.10 (0.04 to 0.16)[c]
Second opinion (no)	0.46 (0.00 to 0.92) [a]	0.10 (0.00 to 0.20) [a]	0.32 (-0.14 to 0.77)	0.07 (-0.03 to 0.17)
Duration of symptoms in months	-0.01 (-0.01 to -0.00)[c]	-0.05 (-0.06 to -0.03)[c]	-0.01 (-0.01 to -0.00)[c]	-0.05 (-0.06 to -0.03)[c]
Dominant side (left)				
Right	-0.06 (-0.31 to 0.20)	-0.01 (-0.07 to 0.04)	-0.05 (-0.30 to 0.20)	-0.01 (-0.07 to 0.04)
Ambidextrous	-0.25 (-0.72 to 0.21)	-0.06 (-0.16 to 0.05)	-0.21 (-0.67 to 0.26)	-0.05 (-0.15 to 0.06)
Recurrence (yes)	-0.62 (-0.88 to -0.36)[c]	-0.14 (-0.20 to -0.08)[c]	-0.61 (-0.86 to -0.35)[c]	-0.13 (-0.19 to -0,08)[c]
PROMs for physical and mental health				
VAS function			0.01 (0.00 to 0.01)[c]	0.04 (0.02 to 0.06)[c]
VAS pain			0.00 (-0.00 to 0.01)	0.01 (-0.01 to 0.03)
EQ5D index score			1.72 (1.25 to 2.18)[c]	0.08 (0.06 to 0.10)[c]
PCS pain catastrophizing score			-0.02 (-0.03 to -0.01)[c]	-0.04 (-0.06 to -0.02)[c]
PHQ depression score			-0.07 (-0.16 to 0.01)	-0.02 (-0.04 to 0.00)
PHQ anxiety score			-0.04 (-0.12 to 0.03)	-0.01 (-0.03 to 0.01)
Illness perception				
B-IPQ consequences				
B-IPQ timeline				
B-IPQ personal control				
B-IPQ identity				
B-IPQ concern				
B-IPQ coherence				

Model 3		Most-definitive model	
B (95% CI)	**β (95% CI)**	**B (95% CI)**	**β (95% CI)**
0.02 (0.01 to 0.03)[c]	0.07 (0.05 to 0.09)[c]	0.02 (0.01 to 0.03)[c]	0.07 (0.05 to 0.09)[c]
-0.25 (-0.40 to -0.10)[c]	-0.06 (-0.09 to -0.02)[c]	-0.24 (-0.39 to -0.09)[b]	-0.05 (-0.09 to -0.02)[b]
0.28 (0.11 to 0.46)[b]	0.06 (0.02 to 0.10)[b]	0.29 (0.11 to 0.46)[b]	0.06 (0.02 to 0.10)[b]
0.26 (0.07 to 0.45)[b]	0.06 (0.02 to 0.10)[b]	0.26 (0.07 to 0.44)[b]	0.06 (0.02 to 0.10)[b]
0.29 (0.04 to 0.55)[a]	0.06 (0.01 to 0.12)[a]	0.29 (0.04 to 0.55)[a]	0.06 (0.01 to 0.12)[a]
0.24 (-0.20 to 0.67)	0.05 (-0.04 to 0.15)	0.28 (-0.16 to 0.71)	0.06 (-0.04 to 0.16)
-0.00 (-0.01 to -0.00)[c]	-0.03 (-0.05 to -0.02)[c]	-0.00 (-0.01 to -0.00)[c]	-0.03 (-0.05 to -0.02)[c]
-0.05 (-0.29 to 0.19)	-0.01 (-0.06 to 0.04)	-0.05 (-0.28 to 0.19)	-0.01 (-0.06 to 0.04)
-0.15 (-0.58 to 0.30)	-0.03 (-0.13 to 0.07)	-0.13 (-0.57 to 0.31)	-0.03 (-0.13 to 0.07)
-0.36 (-0.61 to -0.12)[b]	-0.08 (-0.13 to -0.03)[b]	-0.38 (-0.62 to -0.13)[b]	-0.08 (-0.14 to -0.03)[b]
0.01 (0.01 to 0.01)[c]	0.05 (0.03 to 0.06)[c]	0.01 (0.01 to 0.01)[c]	0.05 (0.03 to 0.06)[c]
0.00 (-0.00 to 0.01)	0.01 (-0.01 to 0.03)	0.00 (-0.00 to 0.01)	0.01 (-0.01 to 0.03)
1.61 (1.16 to 2.06)[c]	0.07 (0.05 to 0.09)[c]	1.63 (1.18 to 2.08)[c]	0.07 (0.05 to 0.09)[c]
0.01 (0.00 to 0.02)[a]	0.02 (0.00 to 0.04)[a]	0.01 (0.00 to 0.02)[a]	0.02 (0.00 to 0.04)[a]
-0.06 (-0.14 to 0.03)	-0.01 (-0.04 to 0.01)	-0.06 (-0.14 to 0.03)	-0.01 (-0.04 to 0.01)
0.04 (-0.03 to 0.11)	0.01 (-0.01 to 0.03)	0.04 (-0.03 to 0.11)	0.01 (-0.01 to 0.03)
0.16 (0.12 to 0.19)[c]	0.09 (0.07 to 0.11)[c]	0.16 (0.12 to 0.19)[c]	0.09 (0.07 to 0.11)[c]
-0.37 (-0.40 to -0.34)[c]	-0.23 (-0.24 to -0.21)[c]	-0.37 (-0.40 to -0.34)[c]	-0.23 (-0.24 to -0.21)[c]
0.01 (-0.02 to 0.04)	0.00 (-0.01 to 0.02)	0.01 (-0.02 to 0.04)	0.01 (-0.01 to 0.02)
0.08 (0.05 to 0.11)[c]	0.05 (0.03 to 0.07)[c]	0.08 (0.05 to 0.11)[c]	0.05 (0.03 to 0.06)[c]
-0.13 (-0.16 to -0.09)[c]	-0.08 (-0.10 to -0.06)[c]	-0.13 (-0.16 to -0.10)[c]	-0.08 (-0.10 to -0.06)[c]
0.24 (0.21 to 0.28)[c]	0.12 (0.10 to 0.13)[c]	0.24 (0.21 to 0.27)[c]	0.12 (0.10 to 0.13)[c]

Supplemental Table 5. *(continued)*

	Model 1		Model 2	
Explanatory variables	**B (95% CI)**	**β (95% CI)**	**B (95% CI)**	**β (95% CI)**
B-IPQ emotional response				
Treatment characteristics				
Type of treatment (minor surgery)				
Type of treatment (major surgery)				
Mean improvement pain				
Mean improvement function				
Multilevel partitioning r^2	0.01		0.03	

In each additional model, more variables potentially explaining expectations are included. Both the unstandardized estimates (B) and standardized estimates (β) are reported with 95% CIs. The nonstandardized estimates (B) in our most-definitive model indicate that with every unit increase in a continuous, dichotomous, or categorical independent variable, the outcome expectations increase or decrease by the value of the nonstandardized estimate (B); standardized estimates (β) are converted to the same scale, which makes it easier to make between-variable comparisons and determine the relative association of each explanatory variable. [a]$p \leq 0.05$; [b]$p \leq 0.01$; [c]$p \leq 0.001$. EQ-5D = EuroQol-5 Dimensions; B-IPQ = Brief Illness Perception Questionnaire; PCS = Pain Catastrophizing Scale; PHQ = Patient Health Questionnaire.

	Model 3		Most-definitive model	
	B (95% CI)	**β (95% CI)**	**B (95% CI)**	**β (95% CI)**
	-0.06 (-0.09 to -0.03)[c]	-0.04 (-0.06 to -0.02)[c]	-0.06 (-0.09 to -0.03)[c]	-0.04 (-0.06 to -0.02)[c]
			2.53 (1.97 to 3.10)[c]	0.56 (0.44 to 0.68)[c]
			2.75 (2.21 to 3.29)[c]	0.61 (0.49 to 0.73)[c]
			-0.02 (-0.05 to 0.01)	-0.05 (-0.11 to 0.02)
			0.11 (0.07 to 0.15)[c]	0.17 (0.11 to 0.24)[c]
	0.12		0.29	

5

DO SELF-REPORTED ILLNESS PERCEPTIONS, PAIN CATASTROPHIZING, AND PSYCHOLOGICAL DISTRESS CHANGE FOLLOWING HAND SURGEON CONSULTATION? A PROSPECTIVE COHORT STUDY

de Ridder WA, van der Oest MJW, Slijper HP, Vermeulen GM, Hovius SER, Selles RW, Hand-Wrist Study Group, Wouters RM;
2023; Journal of Psychosomatic Research

Abstract

Background

Baseline mindset factors are important factors that influence treatment decisions and outcomes. Theoretically, improving the mindset prior to treatment may improve treatment decisions and outcomes. This prospective cohort study evaluated changes in patients' mindset following hand surgeon consultation. Additionally, we assessed if the change in illness perception differed between surgical and nonsurgical patients.

Methods

The primary outcome was illness perception, measured using the total score of the Brief Illness Perception Questionnaire (B-IPQ, range 0-80). Secondary outcomes were the B-IPQ subscales, pain catastrophizing (measured using the Pain Catastrophizing Scale (PCS)), and psychological distress (measured using the Patient Health Questionnaire-4).

Results

A total of 276 patients with various hand and wrist conditions completed the mindset questionnaires before and after hand surgeon consultation (median time interval: 15 days). The B-IPQ total score improved from 39.7 (±10.6) before to 35.8 (±11.3) after consultation ($p<0.0001$, Cohen's $d=0.36$); scores also improved for the B-IPQ subscales Coherence, Concern, Emotional Response, Timeline, Treatment Control, and Identity and the PCS. There were no changes in the other outcomes. Surgical patients improved on the B-IPQ subscales Treatment Control and Timeline, while nonsurgical patients did not.

Conclusions

Illness perception and pain catastrophizing improved following hand surgeon consultation, suggesting that clinicians may actively influence the patients' mindset during consultations, and that they may try to enhance this effect to improve outcomes. Furthermore, surgical patients improved more in illness perceptions, indicating that nonsurgical patients may benefit from a more targeted strategy for changing mindset.

Introduction

Patients visit a clinician to improve their health, which clinicians aim to do, for example, by providing treatments that reduce pain, improve function, or restore form. Common treatments to achieve these goals in musculoskeletal healthcare may include surgery or therapy. We know that the patients' interpretation of their illness, or illness perception, greatly impacts the outcomes of such treatments, i.a., in patients with common musculoskeletal pathology[1-8]. People need to make sense of what and why certain things happen to them, including illness. Patients develop cognitive and emotional illness perceptions that affect how they cope with their health issues. These perceptions may include beliefs about the illness, for example, the meaning, cause, and consequences, the ability to control the illness or recovery, or how long the illness will last[9]. Illness perceptions form part of the self-regulation model formulated by Leventhal et al[9]. They are believed to be involved in a continuous feedback loop, where the illness triggers a particular perception. The patient's coping mechanisms can mediate this perception and either enhance or repress it, leading to various health outcomes[9]. By modifying illness perceptions early on, individuals can adopt more adaptive coping responses and reduce the perceived threat of illness, anxiety, and distress.

Illness perception is part of the patient's mindset, which can be defined as the set of attitudes held by someone, where attitudes include a way of thinking or feeling about someone or something reflected in a person's behavior[10]. The association of other aspects of the patient's mindset with outcomes has also been assessed. In patients with hand and wrist conditions, these include associations of pain catastrophizing and psychological distress with worse pain, slower return to work, less improvement in functional outcomes, and worse satisfaction with treatment results[6,11-16]. Additionally, patients scheduled for surgical treatment have a worse mindset compared to their nonsurgical counterparts, suggesting that treatment decisions also depend on psychological aspects[17].

Van der Oest et al. showed that a positive change in mindset during nonsurgical treatment for thumb base osteoarthritis was associated with more pain reduction during the first three months of treatment[18]. Also, Teuns et al. observed that effective coaching during surgeon consultation improved the range of motion in patients with an isolated minimally displaced fracture of the radial head, suggesting that coaching limited the counterproductive influence of catastrophic thinking[19]. Moreover, psychological interventions (e.g., psycho-education, mindfulness, etc.) successfully improve the patients' mindset, and they may thereby improve outcomes[20-23].

Since positively changing the mindset has several benefits, it is important to investigate if clinicians can influence these psychological factors actively, as, even in a first consult, some events occur that may influence these mindset factors. Clinicians provide information, explain, give answers, and discuss symptoms, concerns, and treatments, in

theory, all affecting the patient's mindset. It is currently unknown whether these events during surgeon consultations indeed influence mindset factors and, if so, to what extent. Although clinicians in musculoskeletal healthcare increasingly acknowledge that mindset is important, they may find it difficult to see how they can positively influence patient mindset. Theoretically, affecting the patient's mindset before treatment, i.e., during the first hand surgeon consultation, may improve treatment decisions and outcomes. Moreover, if surgeons can influence the patient's mindset during their consultation, this would be a relatively easy and practical way to improve treatment decisions and outcomes.

If the patient's mindset does change following consultation, this change may differ between patients scheduled for surgery. We previously observed associations between a worse patient mindset and being scheduled for surgical treatment[17]. As preventing surgery (e.g., in chronic conditions such as thumb base osteoarthritis) can have several benefits, such as fewer complications, lower costs, and shorter rehabilitation, it is important to investigate if the possible effects of the consultation differ between those who eventually undergo surgery and those who do not. In addition, previous studies[17,24] have also shown that patients undergoing nonsurgical treatment have less positive treatment outcome expectations and lower treatment control, meaning that they expect less effect of their treatment compared to patients undergoing surgical treatment. As treatment outcome expectations and treatment control contribute largely to treatment outcomes, it is important to investigate if the possible effects of the consultation are different for surgical and nonsurgical treatment. Especially nonsurgical treatment may benefit from more positive outcome expectations and treatment control to indirectly improve treatment outcomes.

The aim of this study was to assess the magnitude of the change in illness perception after a hand surgeon consultation in a generic group of patients with hand or wrist conditions. Secondary outcomes were subdomains of illness perception (such as personal control over illness, concern, and the coherence of the illness), psychological distress, and pain catastrophizing. In addition, we evaluated whether any change in the outcomes differed between patients scheduled for nonsurgical or surgical treatment.

Methods

Study design

This was a prospective cohort study in which we collected patient-reported measures of mindset factors before and after hand surgeon consultation. Data collection was part of usual care, but additional prospective measurements were added for this study. All patients who presented at our clinic for any hand condition were approached to be included. The local Medical Ethical Committee approved this study prior to data collection, and all patients provided informed consent for the anonymous use of their data. The

study is reported following the STrengthening the Reporting of Observational Studies in Epidemiology statement (STROBE)[25].

Setting

Data were collected at Xpert Clinics, currently comprising 25 hand and wrist surgery clinics and therapy clinics in The Netherlands. Xpert Clinics employs twenty-eight specialized hand surgeons and over 200 hand therapists. At Xpert Clinics, patients can receive insurance-covered care without any barriers or limitations imposed by public health insurance since insurance is mandatory in the Netherlands. Between March 2021 and May 2021, for this prospective study, we invited all new patients at registration and before surgeon consultation to complete a set of mindset questionnaires. This was a regular surgeon consultation, including history, diagnostics, information provision about the diagnosis and treatment options, and a shared-decision making process for drafting a treatment plan (if applicable).

If surgical or nonsurgical hand therapy treatment (hereinafter referred to as treatment) was initiated following the hand surgeon consultation, a hand therapist assigned a measurement track, and routinely collected patient-reported outcome measurements are emailed to the patient, with the exception that no data were collected in patients receiving steroid injections or patients with a “wait and see” policy. Also, we had a lower clinician-compliance in including patients receiving nonsurgical treatment[26], resulting in an overrepresentation of surgical treatment in this study. The measurement track contained a predefined set of measurements employed at predefined time points for selected patient populations[27]. The hand therapist that assigned the measurement track also provides additional information about the diagnosis and the proposed treatment. We did not intervene in the content of the surgeon consultation or the associated session with the hand therapist. That is, these sessions took place during usual care, and there may have been a heterogeneity in the exact fulfillment of these contact moments due to the variety that characterizes the daily practice of an outpatient clinic.

Whereas the measures prior to the surgeon consultation were prospectively collected uniquely for this study, the mindset questionnaires distributed after the consultation were part of the routine outcome measurements. Thus, patients in this study were invited to complete the mindset questionnaires again after surgeon consultation but before the actual start of the treatment; thus, no patient underwent treatment prior to completing the mindset questionnaires again. The median time between completing the questionnaires before and after the consultation was 15 days (Inter Quartile Range: 8-23). This range is similar to ranges used in test-retest reliability studies[28], and the concepts under study are unlikely to change based on a time effect within this time range.

Data were collected using GemsTracker electronic data capture tools (GemsTracker 2020, Erasmus MC and Equipe Zorgbedrijven, Rotterdam/Eindhoven, The Netherlands).

GemsTracker is a secure internet-based application for distributing questionnaires and forms during clinical research and quality registrations. More details on our routine outcome measurement system and procedure are described elsewhere[27].

Participants

Participants were eligible for analysis if they were adults who completed the mindset questionnaires both before and after surgeon consultation. Because this study aimed to assess the magnitude of change in mindset between before and after the first hand surgeon consultation in a general population of patients with hand or wrist complaints, we included patients from all measurement tracks. There were no additional exclusion criteria.

To assess potential selection bias due to non-response, we compared responder and non-responder demographics. Non-responders were defined as patients who did not complete the mindset questionnaires after the first consultation. Responders were defined as patients who completed the mindset questionnaires both before and after the consultation. Apart from a small difference in age between responders and non-responders, we found no other differences between these groups (Supplementary Table 1). Additionally, a non-significant Little's test ($p = 0.25$) further suggested that the data were missing completely at random[29-31].

Variables and measurement

The primary outcome was the change in illness perception following hand surgeon consultation, measured using the total score of the Brief Illness Perception Questionnaire (B-IPQ, range 0-80, lower scores indicate more positive illness perception). The B-IPQ is a reliable and valid tool[32-34] to briefly measure how patients perceive their illness across eight different domains. These domains were secondary outcomes. Each domain is assessed with a single question and answered on an 11-point scale. Higher scores indicate more negative illness perceptions for the questions on how much the patient's illness affects their life (Consequences), how long they expect their illness will last (Timeline), how much they experience symptoms due to their illness (Identity), how concerned they are about their illness (Concern), and how much their illness affects them emotionally (Emotional Response). Higher scores indicate more positive illness perceptions for the remaining questions: the degree of control patients feel they have over their illness (Personal Control), the extent to which the patient think the treatment will help (Treatment Control), and how well they understand their illness (Coherence)[32].

Our secondary outcomes were the B-IPQ subscales, psychological distress, and pain catastrophizing. We measured psychological distress using the 4-item Patient Health Questionnaire (PHQ-4), a valid and reliable tool. This questionnaire consists of two questions about depression and two about anxiety (score range: 0 (no psychological distress) to 12 (severe psychological distress))[35]. Pain catastrophizing was measured

using the 13-item Pain Catastrophizing Scale (PCS), which is valid and reliable. This questionnaire contains questions on rumination, magnification, and helplessness with respect to pain (score range 0-52, higher scores indicate more catastrophizing)[36]. We used validated Dutch translations of these questionnaires[33,37,38].

Sociodemographic characteristics collected at baseline included age, sex, measurement track, duration of symptoms, occupational status (unemployed or light, medium, or heavy physical labor), whether the dominant hand was treated, and whether it concerned a second opinion.

Sample size

To answer our research questions, an a-priori power analysis for paired t-tests showed that a group of 210 patients was needed to demonstrate an effect size of 0.25 with an alpha of 0.05 and a power of 0.95. Our sample of 276 was thus more than sufficient.

Statistical analysis

We used paired t-tests to investigate differences between the pre and post-surgeon consultation measures of illness perception, psychological distress, and pain catastrophizing. Secondly, we performed a stratified analysis for surgical and nonsurgical patients. Primarily, there was a risk of confounding by indication as the surgical and nonsurgical patients may have differed on other aspects, which in turn may have influenced the outcomes. However, by correcting for the pre-consultation psychological profile scores, as we have done, we ensured that the remaining effects could be attributed to the group allocation itself. We used a linear mixed model to assess whether patients receiving surgical or nonsurgical treatment had a different illness perception, psychological distress, or pain catastrophizing after consultation. In these models, we corrected for the pre-consultation scores and used treatment type as a fixed factor. Also, we investigated within-group differences in this subgroup analysis using paired t-tests. All analyses were performed using R Statistical Programming, version 3.3.4 (R Project for Statistical Computing). Cohen's d was used as a measure for the effect size (0.2 = small effect size; 0.5 = medium effect size; 0.8 = large effect size)[39]. A p-value <0.05 was considered statistically significant. However, as calculating a high number of p-values raises a multiple testing issue, the analyses of our secondary outcomes should be considered exploratory.

Results

Participants

Treatment was initiated for 380 patients that had completed the mindset questionnaires before the surgeon consultation. Finally, 276 patients completed the mindset questionnaires both before and after surgeon consultation, leaving 73% (276/380) for inclusion in the final sample (Fig. 1). Seventy-one percent (197/276) of the patients were scheduled for surgery, the other 29% (79/276) for nonsurgical treatment. The baseline characteristics of the included patients show an average representation of our population (Table 1).

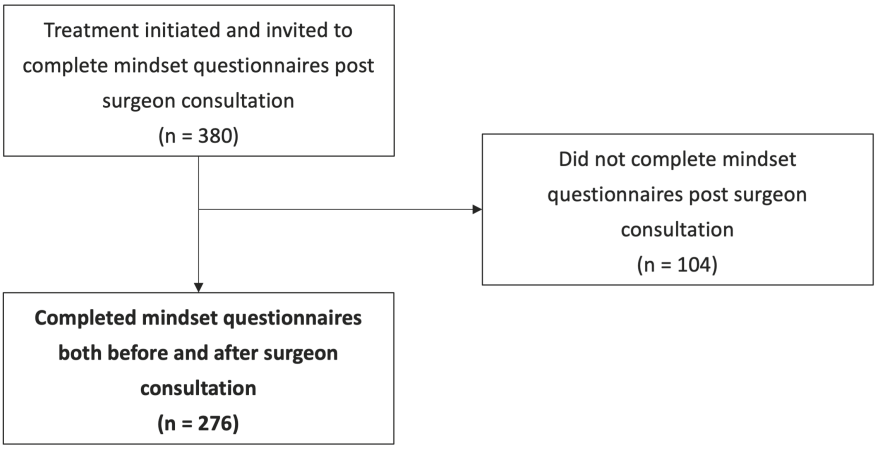


Figure 1. This flowchart illustrates the patient selection for this study.

Table 1. Sociodemographic characteristics at baseline of the patients scheduled for surgery (n = 197), patients scheduled for nonsurgical treatment (n = 79), and all included patients (n = 276)

Variable	Scheduled for surgical treatment (n = 197)	Scheduled for nonsurgical treatment (n = 79)	All patients (n = 276)
Age in years, mean (±SD)	58 (13)	57 (13)	58 ± 13
Male sex, % (n)	36 (71)	28 (22)	34 (93)
Not coming for a second opinion, % (n)	95 (188)	92 (73)	95 (261)
Workload, % (n)			
Unemployed	46 (90)	49 (39)	47 (129)
Light	21 (42)	23 (18)	22 (60)
Medium	24 (47)	23 (18)	24 (65)
Heavy	9 (18)	5 (4)	8 (22)
Symptom duration in months, median (interquartile range)	19.40 (34.92)	10.99 (17.46)	17 ± 31

Table 1. *(continued)*

Variable	Scheduled for surgical treatment (n = 197)	Scheduled for nonsurgical treatment (n = 79)	All patients (n = 276)
Hand dominance, % (n)			
Left	8 (16)	10 (8)	9 (24)
Right	82 (162)	84 (66)	83 (228)
Both	10 (19)	6 (5)	9 (24)
No treatment of the dominant hand, % (n)	47 (93)	65 (51)	52 (144)
Measurement track, % (n)			
Thumb Regular	9 (17)	48 (38)	20 (55)
Thumb Extended	10 (19)	-	7 (19)
Dupuytren	15 (30)	-	11 (30)
Wrist Regular	16 (32)	29 (23)	20 (55)
Wrist Extended	7 (13)	-	5 (13)
Finger Regular	17 (33)	17 (13)	17 (46)
Finger Extended	3 (6)	-	2 (6)
Nerve	24 (47)	6 (5)	19 (52)
Days between before and after consultation, median (interquartile range)	15 (9, 24]	13 (6, 19)	15 (8, 23)

Change in illness perception, pain catastrophizing, and psychological distress

The B-IPQ total score improved from 39.7 before consultation to 35.8 after consultation ($p<0.001$, Cohen's $d=0.36$).

For the B-IPQ subscales (arranged from the largest to the smallest Cohen's d), there was an improvement in the subscales Coherence from 6.8 before consultation to 7.7 after consultation ($p<0.001$, Cohen's $d=0.35$), Concern from 6.0 to 5.1 ($p<0.001$, Cohen's $d=0.32$), Emotional Response from 4.2 to 3.5 ($p<0.001$, Cohen's $d=0.23$), Timeline from 6.2 to 5.6 ($p<0.001$, Cohen's $d=0.21$), Treatment Control from 7.6 to 7.9 ($p=0.01$, Cohen's $d=0.16$), and Identity (from 5.8 to 5.4 ($p=0.01$, Cohen's $d=0.15$). Furthermore, we observed an improvement in the Pain Catastrophizing Scale from 13.5 before consultation to 11.8 after consultation ($p<0.001$, Cohen's $d=0.17$). We observed no changes in our other secondary outcomes (Table 2).

Table 2. Change in mindset questionnaires completed before and after the first surgeon consultation (n=276) with corresponding p-value and Cohen's d.

Questionnaire	Possible score range	Pre consultation	Post consultation	p-value	Cohen's d
B-IPQ total score	0-80 (high score = worse)	39.7 (10.6)	35.8 (11.3)	<0.001	0.36
B-IPQ Consequences (i.e., How much does your illness affect your life?)	0-10 (high score = more consequences)	6.1 (2.4)	6.3 (2.5)	0.24	0.08
B-IPQ Timeline (i.e., How long do you think your illness will continue?_	0-10 (high score = longer perceived timeline)	6.2 (2.8)	5.6 (2.8)	<0.001	0.21
B-IPQ Personal Control (i.e., How much control do you feel you have over your illness?)	0-10 (high score = more personal control)	4.2 (2.3)	4.5 (2.5)	0.22	0.12
B-IPQ Treatment Control (i.e., How much do you think your treatment can help your illness?)	0-10 (high score = more treatment control)	7.6 (1.9)	7.9 (1.8)	0.01	0.16
B-IPQ Identity (i.e., How much do you experience symptoms from your illness?)	0-10 (high score = many severe symptoms)	5.8 (2.4)	5.4 (2.8)	0.01	0.15
B-IPQ Concern (i.e., How concerned are you about your illness?)	0-10 (high score = very concerned)	6.0 (2.7)	5.1 (2.9)	<0.001	0.32
B-IPQ Coherence (i.e., How well do you feel you understand your illness?)	0-10 (high score = better understanding)	6.8 (2.6)	7.7 (2.5)	<0.001	0.35
B-IPQ Emotional Response (i.e., How much does your illness affect you emotionally? (e.g., does it make you angry, scared, upset or depressed?)	0-10 (high score = worse)	4.2 (3.0)	3.5 (3.0)	<0.001	0.23
Patient Health Questionnaire	0-4 (high score = worse)	1.6 (2.5)	1.5 (2.4)	0.32	0.04
Pain Catastrophizing Scale	0-52 (high score = worse)	13.5 (9.8)	11.8 (10.1)	<0.001	0.17

Scores reflect mean (SD) values.

Abbreviations: SD = Standard Deviation, B-IPQ = Brief Illness Perception Questionnaire

Interpretation of Cohen's d: >0.2 is a small effect size; >0.5 is a medium effect size; >0.8 is a large effect size.

Differences between patients scheduled for nonsurgical or surgical treatment

We found differences between patients scheduled for nonsurgical or surgical treatment in the B-IPQ subscales Timeline and Treatment Control (Fig. 2A and 2B). For B-IPQ Treatment Control, we found between-group differences between nonsurgical and surgical patients both before (7.1 ± 1.8 versus 7.9 ± 1.8, p = 0.001) and after consultation (6.9 ± 1.8 versus 8.5 ± 1.5, p<0.001) (a high score indicates more treatment control). There was no within-group change in patients scheduled for nonsurgical treatment (mean difference: -0.2 (95% Confidence Interval: -0.6 to 0.3), p = 0.48), but there was in patients that were eventually scheduled for surgery (mean difference: 0.5 (0.2 to 0.8), p < 0.001).

For B-IPQ Timeline, there was no between-group difference between nonsurgical and surgical patients before consultation. Still, there was a between-group difference post-consultation (6.5 ± 2.7 versus 5.2 ± 2.7, p = 0.001) (a high score indicates a longer perceived timeline). We found no within-group change in nonsurgical patients (mean difference: 0 (-0.5 to 0.4), p = 0.88), but there was a within-group change in the surgical patients (mean difference -0.7 (-1.1 to -0.4), p<0.001). We observed no other differences between the two groups.

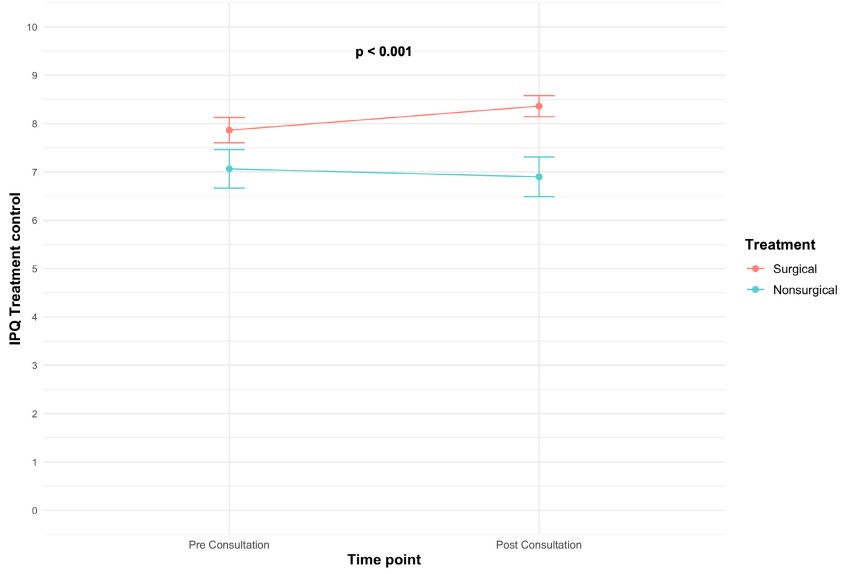


Figure 2A. This figure shows the difference between patients scheduled for nonsurgical (in blue) and surgical (in red) treatment on the Brief Illness Perception Questionnaire subscales (range: 0-10) Treatment Control (high score = more treatment control) (A) and Timeline (high score = longer perceived timeline) (B).

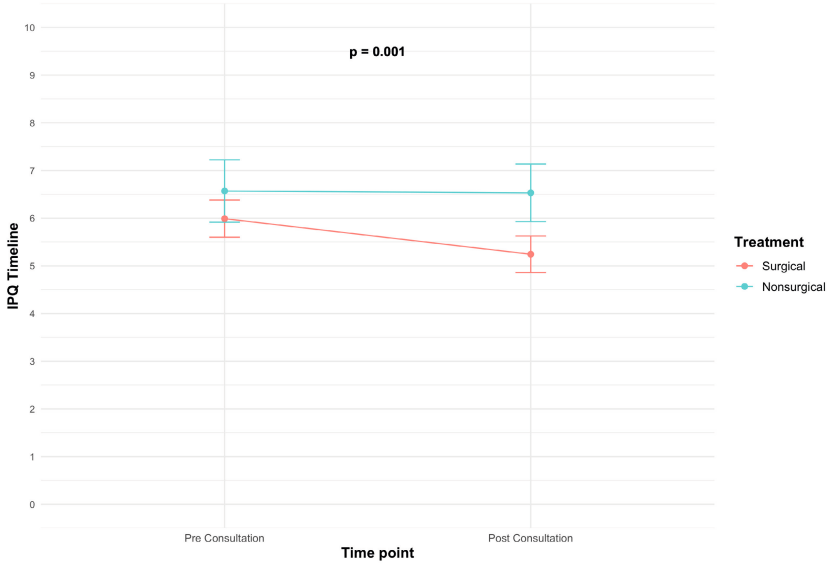


Figure 2B. *(continued)*

Discussion

We found an improvement in overall illness perception, changes in six of the eight subscales of illness perception, and an improvement in pain catastrophizing. Additionally, we found a greater change in treatment control and perceived timeline in surgical patients than in nonsurgical patients. Our study indicates that hand surgeons may influence the patients' illness perception and pain catastrophizing during consultations. They might try to enhance this effect to improve treatment outcomes further. Although we cannot assume causality, this seems a relatively easy and practical way to improve treatment decisions and outcomes. Furthermore, the differences observed between surgical and nonsurgical patients indicate that hand surgeons may not fully exploit the potential of nonsurgical treatment.

Change in illness perception, pain catastrophizing, and psychological distress

We found improvements in our primary and almost all secondary outcomes after one single surgeon consultation, although these secondary outcomes should be considered exploratory. To our knowledge, no other study investigated the influence of the first surgeon consultation on patient mindset, so we cannot compare our findings with previous work. However, given that illness perceptions are shaped by former experience, interpretation of symptoms, and information provision[9], our results are not unexpected or surprising. As the initial consultation with a surgeon adds to this former experience, helps patients to interpret their symptoms, and provides information on their condition. Thus, one may have also considered it strange if the improvements we found would not

take place. In addition to the particular context with a surgeon in a white coat (which may induce a positive placebo effect), this information might also help to comfort the patient, decreasing concern and, possibly, the tendency to catastrophize pain. A noteworthy finding is that we found changes in pain catastrophizing, indicating that it is a dynamic instead of a stable trait, which has been debated in literature[40,41]. Moreover, our findings are consistent with studies in other fields, which demonstrated that even small interventions can significantly affect illness perception. For instance, Devcich et al. found that results of a coronary angiography can immediately influence how patients perceive and emotionally respond to their symptoms[42]. Similarly, other brief interventions improved illness perceptions in myocardial infarction patients[23] and their spouses[43]. These examples highlight the potential for small events to influence illness perception, as seen in our study. While we only investigated patients with hand and wrist conditions, our findings may be generalizable to a broader group of patients with musculoskeletal conditions. Future research may investigate this.

Our study provides valuable leads for interventions to improve outcomes, as illness perception and pain catastrophizing are such essential aspects, and we have now shown that it may be influenced relatively easily during only one surgeon consultation. It would be interesting to investigate why there was no change in personal control; future research may aim to influence that aspect of illness perception. Hypothetically, since more personal control reflects an internal locus of control, improving personal control could improve treatment coherence and, thereby, treatment outcomes. Moreover, as previous research found that patients with thumb base osteoarthritis scheduled for surgery have worse personal control than those scheduled for nonsurgical treatment[17], improving personal control may, hypothetically, prevent unnecessary surgical treatment.

One may question whether the change found is clinically relevant and greater than the measurement error. To conclude on this, the changes should be equal to or larger than the Minimal Important Change (MIC) or the Smallest Detectable Change (SDC). To our knowledge no MIC and SDC values for the questionnaires used in our study have been reported. However, SDCs on a group level can be calculated by dividing the SDC on an individual level by $\sqrt{n}$ [44]. Using individual SDC values of previous studies on the B-IPQ total score[45-47], the SDC on a group level for the B-IPQ in our sample of 276 participants would range from 0.13 to 0.91. Since we found an improvement of almost 4 points on the total score, this indicates that we found a real change. Similarly, the changes in our secondary outcomes also extend these group SDCs[33,45-47]. Moreover, we found a small to medium effect size (values ranging from 0.17 to 0.36) on our significant outcomes. Future research should confirm the clinical importance of our findings and may investigate MIC values for the questionnaires we used.

Differences between patients scheduled for nonsurgical or surgical treatment

Our findings of the differences between patients scheduled for surgical treatment versus nonsurgical treatment align with other studies. Several studies showed that patients with diverse hand or wrist conditions scheduled for surgery have higher expectations and more treatment control than patients scheduled for nonsurgical treatment[7,17,24,48]. Our finding that patients scheduled for surgery also have a more positive view following surgeon consultation (i.e., they expect their illness to last shorter and experience more treatment control) confirms findings directing to higher, more positive expectations in patients scheduled for surgery[7,17,48]. This also highlights that nonsurgical treatment has a worse image than surgical treatment, underlining that we may need to focus more on boosting expectations of nonsurgical treatment, as less invasive, nonsurgical treatment options may currently not be fully utilized. Future studies may investigate this in more detail. Possible directions may be discussing the patients' views about their illness and treatment, using decision-support tools such as prediction models, and more extensive or patient-specific education on the illness or treatment, e.g., by using explanatory or interactive videos.

Limitations

A limitation of our observational design is the number of patients that did not respond. Selective non-response could lead to selection bias. However, our non-responder analysis indicated only a small difference in age, and Little's test further suggested that the data were missing completely at random. Therefore, we are confident that this did not influence our results.

Another limitation related to the observational design is that we included a relatively high number of patients that underwent surgery. At the time of this study, no outcomes were collected in patients with steroid injections or a "wait and see" policy. Moreover, the clinician-compliance with initiating measurements in patients undergoing nonsurgical treatments was lower than in those undergoing sugery[26]. As our inclusion depended on this measurement assignment, the number of surgical patients is overrepresented in this study and does not reflect the actual distribution at Xpert Clinics.

Another aspect of our observational design is that we are unsure that the surgeon consultation actually caused the changes in the patient's mindset as we did not use randomization. It is possible that other events than the surgeon consultation took place within the 15-day interval that may have influenced the patient's mindset. However, since we used test-retest reliable and responsive questionnaires, we can assume with confidence that the changes observed reflect a true change. Although theoretically, symptom reduction due to natural recovery may have occurred, and thereby improvement in the patient's mindset may have happened, it is unlikely that this has affected our results to a great extent. Our time interval was only fifteen days, and most elective hand and wrist treatments are for non-acute conditions (e.g., osteoarthritis), so natural recovery is unlikely.

In our study, there was a risk of confounding by indication as the surgical and nonsurgical patients may have differed on other aspects that influence the outcome. By correcting for the pre-consultation scores, we ensured that the remaining effects could be attributed to the group allocation. Still, from a methodological point of view, the best solution would have been random assignment to surgical and nonsurgical treatment. However, randomly assigning patients to either of these groups would be relatively artificial because, in daily practice, this decision is made during the consultation by the surgeon and patient, and not randomly.

We did not intervene in the exact content of the consultations. Therefore, we are unsure what took place during these moments, and the fulfillment of these moments may have varied, e.g., due to clinician factors, patient factors, or other factors that reflect the unruly daily practice of the outpatient clinic. Future studies may investigate a change in illness perception and pain catastrophizing following surgeon consultation in a standardized setting, i.e., by using a randomized design with an experimental group that receives the surgeon consultation and a control group that does not.

Conclusions

We observed an improved illness perception and pain catastrophizing following hand surgeon consultation in patients with hand and wrist conditions, suggesting that hand surgeons can actively influence the patients' mindset during a single consultation. As illness perception and pain catastrophizing influence treatment choices and outcomes, our findings suggest that actively influencing surgeons' consultations may further improve the patients' mindset and thereby treatment choices and outcomes. Furthermore, patients scheduled for surgery improved more on treatment control and timeline after surgeon consultation compared to their nonsurgical counterparts. Taken all together, possible interventions during or directly following the first surgeon consultation may aim at addressing illness perceptions and boosting pretreatment expectations of surgical and especially nonsurgical treatments, e.g., by discussing the patients' views about their illness and treatment, the use of decision-support tools such as prediction models, and more extensive or patient-specific education on the illness or treatment.

Acknowledgments

The authors thank all patients who participated and allowed their data to be anonymously used for the present study. The collaborators of the Hand-Wrist Study Group are: RAM Blomme, BJR Sluijter, DJJC van der Avoort, GJ Halbesma, A Kroeze, J Smit, J Debeij, ET Walbeehm, GM van Couwelaar, GM Vermeulen, JP de Schipper, JFM Temming, JH van Uchelen, HL de Boer, KP de Haas, K Harmsen, OT Zöphel, R Feitz, JS Souer, R Koch, SER Hovius, TM Moojen, X Smit, R Hagen, R van Huis, PY Pennehouat, K Schoneveld, YE van Kooij, RM Wouters, J Veltkamp, A Fink, L Esteban Lopez, WA de Ridder, HP Slijper, RW

Selles, JT Porsius, J Tsehaie, R Poelstra, MC Jansen, MJW van der Oest, L Hoogendam, JS Teunissen, JE Koopman, J Dekker, MHP ter Stege, JM Zuidam, CA van Nieuwenhoven, CA Hundepool, BEPA van der Heijden, JW Colaris, WR Bijlsma.

References

1. Broadbent E, Wilkes C, Koschwanez H, Weinman J, Norton S, Petrie KJ. A systematic review and meta-analysis of the Brief Illness Perception Questionnaire. *Psychol Health.* 2015;30(11):1361-1385.
2. Bethge M, Bartel S, Streibelt M, Lassahn C, Thren K. [Illness perceptions and functioning following total knee and hip arthroplasty]. *Z Orthop Unfall.* 2010;148(4):387-392.
3. Kok LM, Vliet Vlieland TP, Fiocco M, Kaptein AA, Nelissen RG. Musicians' illness perceptions of musculoskeletal complaints. *Clin Rheumatol.* 2013;32(4):487-492.
4. Lindberg MF, Miaskowski C, Rustøen T, Rosseland LA, Paul SM, Cooper BA, Lerdal A. The Impact of Demographic, Clinical, Symptom and Psychological Characteristics on the Trajectories of Acute Postoperative Pain After Total Knee Arthroplasty. *Pain Med.* 2017;18(1):124-139.
5. van der Oest MJW, Poelstra R, Feitz R, Vranceanu AM, Slijper HP, Selles RW, Porsius JT. Illness Perceptions of Patients With First Carpometacarpal Osteoarthritis, Carpal Tunnel Syndrome, Dupuytren Contracture, or Trigger Finger. *J Hand Surg Am.* 2020;45(5):455.e451-455.e458.
6. Blackburn J, van der Oest MJW, Chen NC, Feitz R, Duraku LS, Zuidam JM, Vranceanu AM, Selles RW. Are Patient Expectations and Illness Perception Associated with Patient-reported Outcomes from Surgical Decompression in de Quervain's Tenosynovitis? *Clin Orthop Relat Res.* 2021;479(5):1147-1155.
7. Hoogendam L, van der Oest MJW, Wouters RM, Andrinopoulou ER, Vermeulen GM, Slijper HP, Porsius JT, Selles RW. Patients with higher treatment outcome expectations are more satisfied with the results of nonoperative treatment for thumb base osteoarthritis: a cohort study. *Arch Phys Med Rehabil.* 2021;102(8):1533-1540.
8. van der Oest MJW, Hoogendam L, Wouters RM, Vermeulen GM, Slijper HP, Selles RW, Vranceanu AM, Porsius JT. Associations between positive treatment outcome expectations, illness understanding, and outcomes: a cohort study on non-operative treatment of first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2022;44(19):5487-5494.
9. Leventhal H, Phillips LA, Burns E. The Common-Sense Model of Self-Regulation (CSM): a dynamic framework for understanding illness self-management. *J Behav Med.* 2016;39(6):935-946.
10. OED OED. Oxford english dictionary. *Online edition at< dictionary oed com/entrance dtl.* 2010.
11. Hoogendam L, van der Oest MJW, Tsehaie J, Wouters RM, Vermeulen GM, Slijper HP, Selles RW, Porsius JT. Psychological factors are more strongly associated with pain than radiographic severity in non-invasively treated first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021;43(13):1897-1902.
12. De Ridder WA, Wouters RM, Hoogendam L, Vermeulen GM, Slijper HP, Selles RW. Which Factors Are Associated With Satisfaction With Treatment Results in Patients With Hand and Wrist Conditions? A Large Cohort Analysis. *Clin Orthop Relat Res.* 2022;480(7):1287-1301.
13. Blackburn J, van der Oest MJW, Selles RW, Chen NC, Feitz R, Vranceanu AM, Porsius JT. Which Psychological Variables Are Associated With Pain and Function Before Surgery for de Quervain's Tenosynovitis? A Cross-sectional Study. *Clin Orthop Relat Res.* 2019;477(12):2750-2758.

14. Sun PO, Walbeehm ET, Selles RW, Jansen MC, Slijper HP, Ulrich DJO, Porsius JT. Influence of illness perceptions, psychological distress and pain catastrophizing on self-reported symptom severity and functional status in patients with carpal tunnel syndrome. *J Psychosom Res.* 2019;126:109820.

15. London DA, Stepan JG, Boyer MI, Calfee RP. The impact of depression and pain catastrophization on initial presentation and treatment outcomes for atraumatic hand conditions. *J Bone Joint Surg Am.* 2014;96(10):806-814.

16. Jansen MC, van der Oest MJW, de Haas NP, Selles Ph DR, Zuidam Md Ph DJ. The Influence of Illness Perception and Mental Health on Return to Work After Carpal Tunnel Release Surgery. *J Hand Surg Am.* 2021;46(9):748-757.

17. Wouters RM, Vranceanu AM, Slijper HP, Vermeulen GM, van der Oest MJW, Selles RW, Porsius JT. Patients With Thumb-base Osteoarthritis Scheduled for Surgery Have More Symptoms, Worse Psychological Profile, and Higher Expectations Than Nonsurgical Counterparts: A Large Cohort Analysis. *Clin Orthop Relat Res.* 2019;477(12):2735-2746.

18. van der Oest M, Wouters, RM, Bakhshaie, J, Hoogendam, L, Souer, SJ, Vranceanu, AM, Zuidam, JM, the Hand-Wrist Study Group, Selles, RW. Changes in mindset during treatment influences the outcomes of nonsurgical treatment for thumb OA; a cohort study. *Submitted.* 2022.

19. Teunis T, Thornton ER, Guitton TG, Vranceanu AM, Ring D. Coaching of patients with an isolated minimally displaced fracture of the radial head immediately increases range of motion. *J Hand Ther.* 2016;29(3):314-319.

20. Arat S, De Cock D, Moons P, Vandenberghe J, Westhovens R. Modifiable correlates of illness perceptions in adults with chronic somatic conditions: A systematic review. *Res Nurs Health.* 2018;41(2):173-184.

21. Gehrt TB, Wisbech Carstensen TB, Ørnbøl E, Fink PK, Kasch H, Frostholm L. The role of illness perceptions in predicting outcome after acute whiplash trauma: a multicenter 12-month follow-up study. *Clin J Pain.* 2015;31(1):14-20.

22. Zelber-Sagi S, Bord S, Dror-Lavi G, Smith ML, Towne SD, Jr., Buch A, Webb M, Yeshua H, Nimer A, Shibolet O. Role of illness perception and self-efficacy in lifestyle modification among non-alcoholic fatty liver disease patients. *World J Gastroenterol.* 2017;23(10):1881-1890.

23. Broadbent E, Ellis CJ, Thomas J, Gamble G, Petrie KJ. Can an illness perception intervention reduce illness anxiety in spouses of myocardial infarction patients? A randomized controlled trial. *J Psychosom Res.* 2009;67(1):11-15.

24. de Ridder WA, Hoogendam L, Zeroual F, Slijper HP, Wouters RM, Vermeulen GM, Selles RW, van der Oest MJW. Treatment Invasiveness and Illness Perceptions Are Strongly Associated With Outcome Expectations in Patients Treated for Hand or Wrist Conditions: A Cross-sectional Study. *Clin Orthop Relat Res.* 2023;481(5):994-1005.

25. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, Initiative S. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453-1457.

26. Feitz R, van Kooij YE, Ter Stege MHP, van der Oest MJW, Souer JS, Wouters RM, Slijper HP, Selles RW, Hovius SER. Closing the loop: a 10-year experience with routine outcome measurements to improve treatment in hand surgery. *EFORT open reviews.* 2021;6(6):439-450.

27. Selles RW, Wouters RM, Poelstra R, van der Oest MJW, Porsius JT, Hovius SER, Moojen TM, van Kooij Y, Pennehouat PY, van Huis R, Vermeulen GM, Feitz R, Slijper HP. Routine health outcome measurement: development, design, and implementation of the Hand and Wrist Cohort. *Plast Reconstr Surg.* 2020;146(2):343-354.

28. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press; 2011.

29. de Groot JA, Janssen KJ, Zwinderman AH, Bossuyt PM, Reitsma JB, Moons KG. Correcting for partial verification bias: a comparison of methods. *Ann Epidemiol.* 2011;21(2):139-148.

30. Little RJA. A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association.* 1988;83(404):1198-1202.

31. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, Frangakis C, Hogan JW, Molenberghs G, Murphy SA, Neaton JD, Rotnitzky A, Scharfstein D, Shih WJ, Siegel JP, Stern H. The prevention and treatment of missing data in clinical trials. *N Engl J Med.* 2012;367(14):1355-1360.

32. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res.* 2006;60(6):631-637.

33. de Raaij EJ, Schröder C, Maissan FJ, Pool JJ, Wittink H. Cross-cultural adaptation and measurement properties of the Brief Illness Perception Questionnaire-Dutch Language Version. *Man Ther.* 2012;17(4):330-335.

34. Leysen M, Nijs J, Meeus M, Paul van Wilgen C, Struyf F, Vermandel A, Kuppens K, Roussel NA. Clinimetric properties of illness perception questionnaire revised (IPQ-R) and brief illness perception questionnaire (Brief IPQ) in patients with musculoskeletal disorders: A systematic review. *Man Ther.* 2015;20(1):10-17.

35. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics.* 2009;50(6):613-621.

36. Sullivan MJL BS, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychol Assess.* 1995;7:524-521.

37. Timmermans I, Versteeg H, Meine M, Pedersen SS, Denollet J. Illness perceptions in patients with heart failure and an implantable cardioverter defibrillator: Dimensional structure, validity, and correlates of the brief illness perception questionnaire in Dutch, French and German patients. *J Psychosom Res.* 2017;97:1-8.

38. Mertens VC, Moser A, Verbunt J, Smeets R, Goossens M. Content Validity of the Credibility and Expectancy Questionnaire in a Pain Rehabilitation Setting. *Pain Pract.* 2017;17(7):902-913.

39. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.

40. Wade JB, Riddle DL, Thacker LR. Is pain catastrophizing a stable trait or dynamic state in patients scheduled for knee arthroplasty? *Clin J Pain.* 2012;28(2):122-128.

41. Sturgeon JA, Zautra AJ. State and trait pain catastrophizing and emotional health in rheumatoid arthritis. *Ann Behav Med.* 2013;45(1):69-77.

42. Devcich DA, Ellis CJ, Gamble G, Petrie KJ. Psychological responses to cardiac diagnosis: changes in illness representations immediately following coronary angiography. *J Psychosom Res.* 2008;65(6):553-556.

43. Broadbent E, Ellis CJ, Thomas J, Gamble G, Petrie KJ. Further development of an illness perception intervention for myocardial infarction patients: a randomized controlled trial. *J Psychosom Res.* 2009;67(1):17-23.

44. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.

45. Hallegraeff JM, van der Schans CP, Krijnen WP, de Greef MH. Measurement of acute nonspecific low back pain perception in primary care physical therapy: reliability and validity of the brief illness perception questionnaire. *BMC Musculoskelet Disord.* 2013;14:53.

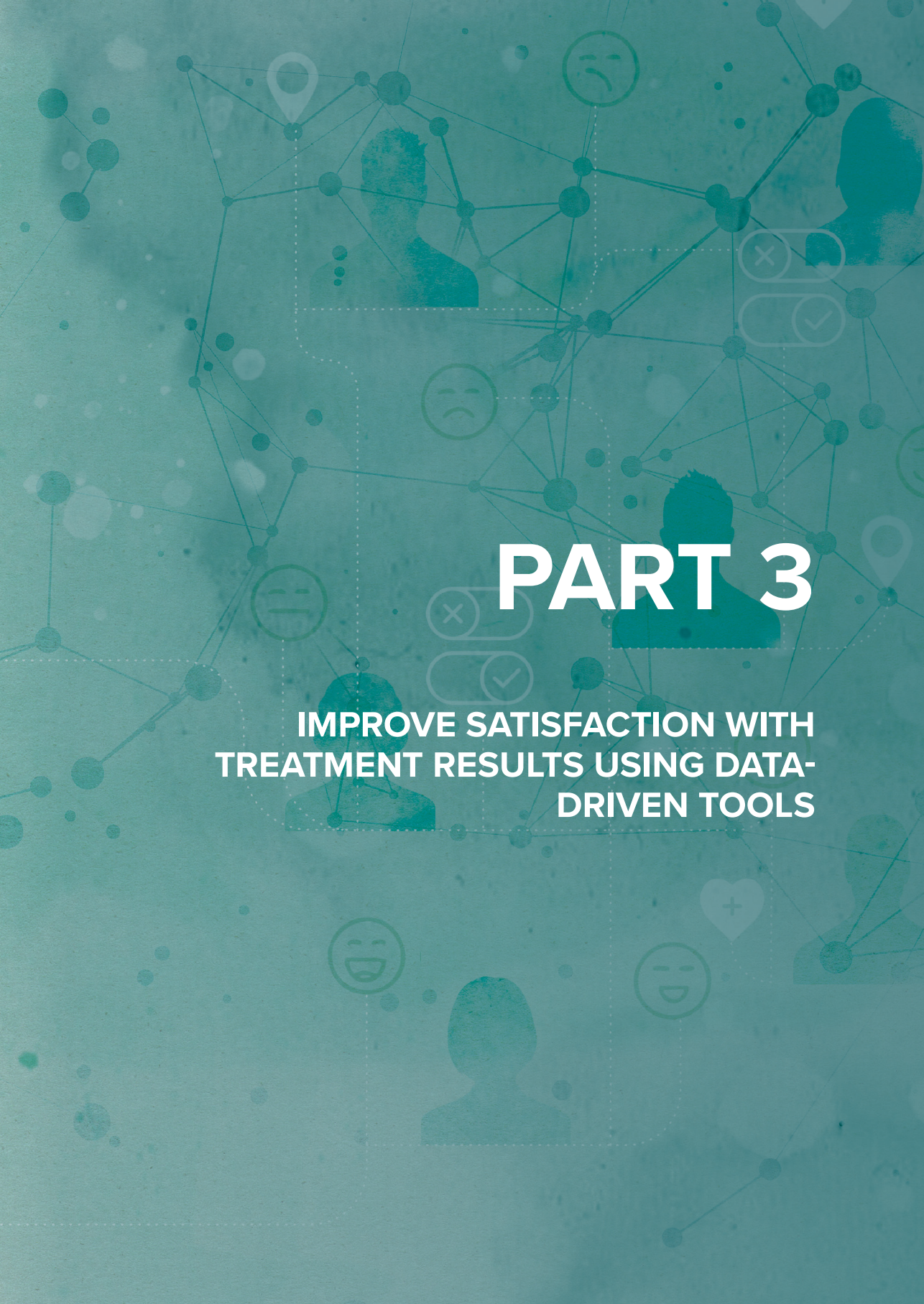
46. Seydi M, Akhbari B, Ghasem Abad SK, Jaberzadeh S, Saeedi A, Ashrafi A, Shakoorianfard MA. Psychometric properties of the Persian version of the brief illness questionnaire in Iranian with non-specific chronic neck pain. *J Bodyw Mov Ther.* 2021;28:323-331.

47. Igwesi-Chidobe CN, Sorinola IO, Godfrey EL. The Igbo Brief Illness Perceptions Questionnaire: A cross-cultural adaptation and validation study in Nigerian populations with chronic low back pain. *J Back Musculoskelet Rehabil.* 2021;34(3):399-411.

48. Hoogendam L, Koopman JE, van Kooij YE, Feitz R, Hundepool CA, Zhou C, Slijper HP, Selles RW, Wouters RM. What Are the Minimally Important Changes of Four Commonly Used Patient-reported Outcome Measures for 36 Hand and Wrist Condition-Treatment Combinations? *Clin Orthop Relat Res.* 2022;480(6):1152-1166.

Supplementary Table 1. Comparison of characteristics for patients who completed all mindset questionnaires both before and after consultation (responders) and patients who did not complete all mindset questionnaires after consultation (non-responders)

	Responders	**Non-responders**	**P value**
	274	**104**	
Age in years	58 ± 13	53 ± 18	**0.008**
Sex (male)	34 (93)	35 (36)	0.96
Second opinion (no)	95 (261)	97 (101)	0.44
Workload			0.65
Unemployed	47 (129)	46 (48)	
Light	22 (60)	23 (24)	
Medium	24 (65)	20 (21)	
Heavy	8 (22)	11 (11)	
Symptom duration in months median (interquartile range)	17 ± 31	19 ± 31	0.63
Hand dominance			0.77
Left	9 (24)	8 (8)	
Right	83 (228)	86 (89)	
Both	9 (24)	7 (7)	
Measurement track name			0.21
Thumb Regular	20 (55)	18 (19)	
Thumb Extended	7 (19)	3 (3)	
Dupuytren	11 (30)	7 (7)	
Wrist Regular	20 (55)	24 (25)	
Wrist Extended	5 (13)	3 (3)	
Finger Regular	17 (46)	25 (26)	
Finger Extended	2 (6)	3 (3)	
Nerve	19 (52)	17 (18)	



PART 3

IMPROVE SATISFACTION WITH TREATMENT RESULTS USING DATA-DRIVEN TOOLS

6

THE ULTRASHORT MENTAL HEALTH SCREENING TOOL IS A VALID AND RELIABLE MEASURE WITH ADDED VALUE TO SUPPORT DECISION-MAKING

Wouters RM, de Ridder WA, Slijper HP, Vermeulen GM, Hovius SER, Selles RW; Hand-Wrist Study Group; van der Oest MJW; 2023; Clinical Orthopaedics and Related Research

Abstract

Background

Mental health influences symptoms, outcomes, and decision-making in musculoskeletal healthcare. Implementing measures of mental health in clinical practice can be challenging. An ultrashort screening tool for mental health with a low burden is currently unavailable but could be used as a conversation starter, expectation management tool, or decision support tool.

Questions/purposes

(1) Which items of the Pain Catastrophizing Scale (PCS), Patient Health Questionnaire (PHQ-4), and Brief Illness Perception Questionnaire (B-IPQ) are the most discriminative and yield a high correlation with the total scores of these questionnaires? (2) What is the construct validity and added clinical value (explained variance for pain and hand function) of an ultrashort four-item mental health screening tool? (3) What is the test-retest reliability of the screening tool? (4) What is the response time for the ultrashort screening tool?

Methods

This was a prospective cohort study. Data collection was part of usual care at Xpert Clinics, The Netherlands, but additional prospective measurements were added for this study. Between September 2017 and January 2022, we included 19,156 patients with hand and wrist conditions. We subdivided these into four samples: a test set to select the screener items (n = 18,034), a validation set to determine whether the selected items were solid (n = 1017), a sample to determine the added clinical value (explained variance for pain and hand function, n = 13,061), and a sample to assess the test-retest reliability (n = 105). Patients were eligible for either sample if they completed all relevant measurements of interest for that particular sample. To create an ultrashort short screening tool that is valid, reliable, and has added value, we began by picking the most discriminatory items (that is, the items that were most influential for determining the total score) from the PCS, PHQ-4, and B-IPQ using chi-square automated interaction detection (a machine learning algorithm). To assess construct validity (how well our screening tool assesses the constructs of interest), we correlated these items with the associated sum score of the full questionnaire in the test set and validation set. We compared the explained variance of linear models for pain and function using the screening tool items or the original sum scores of the PCS, PHQ-4, and B-IPQ to further assess the screening tool's construct validity and added value. We evaluated test-retest reliability by calculating weighted kappas, intraclass correlation coefficients, and the standard error of measurement.

Results

We identified four items and used these in the screening tool. The screening tool items were highly correlated with the PCS (Pearson coefficient = 0.82; $p < 0.001$), PHQ-4 (0.87; $p < 0.001$), and B-IPQ (0.85; $p < 0.001$) sum scores, indicating high construct validity.

The full questionnaires explained only slightly more variance in pain and function (10% to 22%) than the screening tool did (9% to 17%), again indicating high construct validity and much added clinical value of the screening tool. Test-retest reliability was high for the PCS (ICC 0.75, weighted kappa 0.75) and B-IPQ (ICC 0.70 to 0.75, standard error of measurement 1.3 to 1.4) items and moderate for the PHQ-4 item (ICC 0.54, weighted kappa 0.54). The median response time was 43 seconds, against more than 4 minutes for the full questionnaires.

Conclusion

Our ultrashort, valid, and reliable screening tool for pain catastrophizing, psychologic distress, and illness perception can be used before clinician consultation and may serve as a conversation starter, an expectation management tool, or a decision support tool. The clinical utility of the screening tool is that it can indicate that further testing is warranted, guide a clinician when considering a consultation with a mental health specialist, or support a clinician in choosing between more invasive and less invasive treatments. Future studies could investigate how the tool can be used optimally and whether using the screening tool affects daily clinic decisions.

Introduction

In musculoskeletal healthcare, the patient's mental health has gained attention in recent years. Numerous studies have demonstrated that mental health factors influence symptoms, outcomes, and treatment choices [1-18]. For example, patients with thumb-base osteoarthritis scheduled for surgery have worse psychologic profiles than their nonsurgical counterparts [19], suggesting that domains of mental health play an important role in choosing between surgical and nonsurgical treatment. Important mental health domains include pain catastrophizing, psychologic distress (anxiety and depression), and illness perceptions. Given the relevance of mental health in many musculoskeletal conditions, it is valuable to routinely examine one's mental health to support personalized and value-based healthcare and facilitate shared decision-making [20-22].

Several patient-reported measures of mental health are available, including the Pain Catastrophizing Scale (PCS) [23], the four-item Patient Health Questionnaire (PHQ-4) [24], and the Brief Illness Perception Questionnaire (B-IPQ) [25-27], adding up to 25 questions if one would obtain a (relatively) complete picture of a patient's mental health. Implementing these or similar measures in clinical practice can be challenging. Using mental health measures in addition to standard outcome sets (such as for hand and wrist conditions [28]) requires greater time investment from patients and adds to the burden of routine outcome measurements.

Hypothetically, questionnaires with fewer items may yield a higher compliance rate. Another issue of implementing measures of mental health in daily clinical practice is that patients may not understand why they have to complete these questionnaires if, in their opinion, they have very objectifiable symptoms because of a specific physical condition (such as osteoarthritis). Consequently, patients may feel that using these measures to evaluate mental health is inappropriate. Reducing the number of questions while obtaining a valid and reliable picture of a patient's mental health could be a solution. This would also be helpful for clinicians, because many clinicians in musculoskeletal healthcare have little or no time for an in-depth evaluation of mental health during a consultation, and they may also lack the skills for such conversations.

There is a need for a short screening tool that provides an accurate view of patients' mental health with a low patient and clinician burden to overcome these issues. Ideally, such a screening tool would be used before a primary clinician consultation to guide the consultation. A screening tool for mental health would have great clinical relevance because it can be used as a conversation starter, expectation management tool, or decision support tool. For example, it could enable clinicians to discuss the patient's thoughts and feelings and the influence of those thoughts and feelings on perceived symptoms and treatment outcomes, or it may inform the decision to refer a patient to a mental health specialist.

Therefore, we asked: (1) Which items of the PCS, PHQ-4, and B-IPQ are the most discriminative and yield a high correlation with the total scores of these questionnaires? (2) What is the construct validity and added clinical value (explained variance for pain and hand function) of an ultrashort four-item mental health screening tool? (3) What is the test-retest reliability of the screening tool? (4) What is the response time for the ultrashort screening tool?

Patients and Methods

Study Design and Setting

This prospective cohort study followed the STrengthening the Reporting of Observational studies in Epidemiology statement [29]. Data were collected at Xpert Clinics, comprising 25 specialized treatment centers in the Netherlands for hand surgery and therapy. Patient care is reimbursed by Dutch basic insurance. Xpert Clinics currently employs 27 hand surgeons and more than 150 hand therapists. All hand surgeons are certified by the Federation of European Societies for Surgery of the Hand or are fellowship-trained. Data collection was part of usual care, but additional prospective measurements were added for this study. In the routine outcome measurement system, a measurement track is assigned to each patient, including predefined measurements at predefined timepoints. Details on our routine outcome measurement system are described elsewhere [30].

Participants

We used four samples. The first was a test set in which we developed the screening tool and first assessed construct validity (how well our screening tool assesses the constructs of interest). Between September 2017 and January 2022, we treated 37,911 patients for various hand and wrist conditions. Of those, we considered adult patients that completed the mental health measures after clinician consultation as part of their routine outcome measurement as potentially eligible for the test set. These measures were baseline measurements for patients scheduled for either nonsurgical or surgical treatment. Based on that, 48% (18,034) were included in the test set; 52% (19,877) were excluded because of missing data (Fig. 1).

The second sample was a validation set and was used to determine whether the selected items were solid (1017). Between September 2017 and January 2022, we invited an additional 4089 patients with various hand and wrist conditions to complete the mental health measures before consultation. We considered all patients who completed these measures eligible for the validation set. Based on that, 25% (1018) were included in the validation set, and the remaining 75% (3071) were excluded because of missing data.

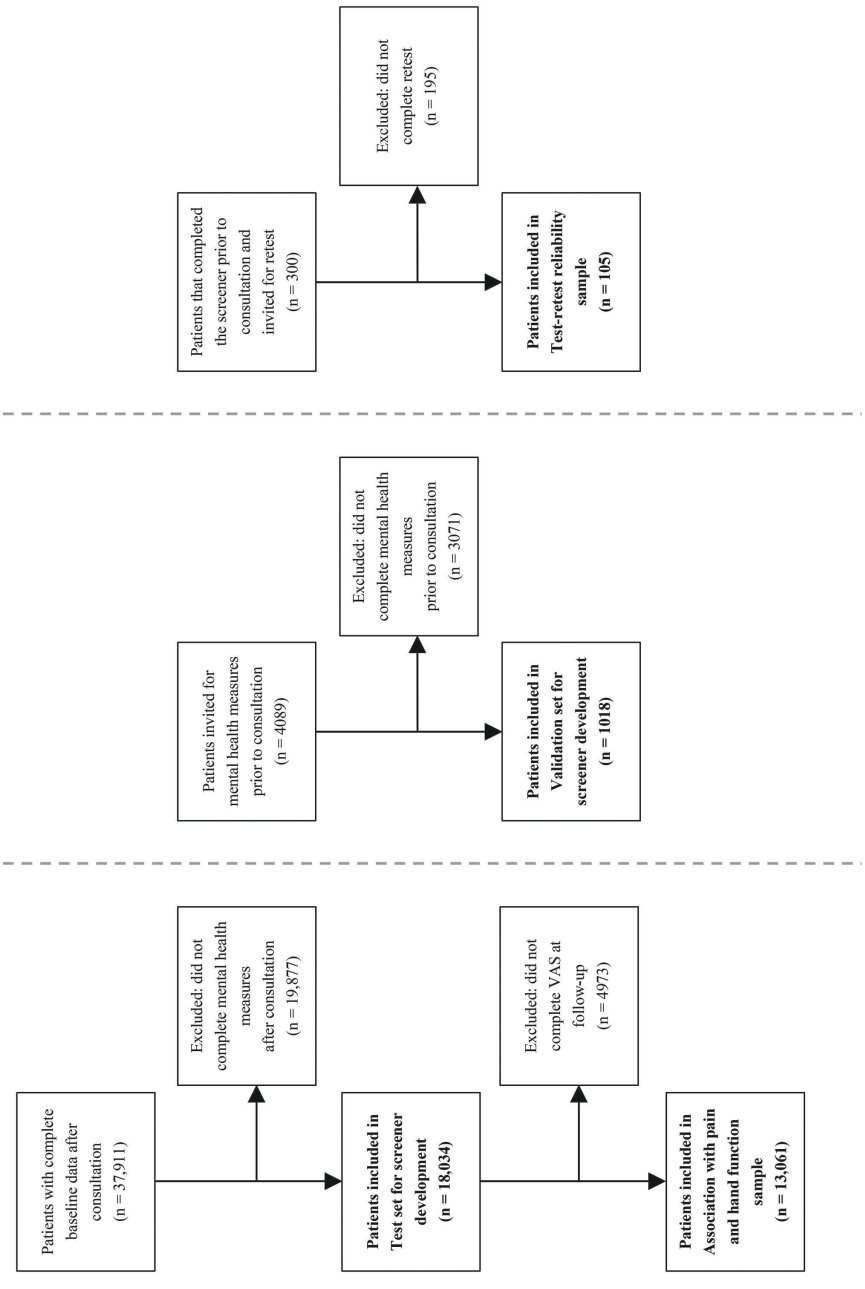


Fig. 1 This flowchart of the study shows the inclusion and exclusion criteria.

To further assess construct validity and added clinical value (explained variance for pain and hand function), we used a third sample to assess the association of the screening tool items with pain and hand function at baseline and at 3 months of follow-up. We considered all patients from the test set who also completed the VAS for pain and function at baseline and 3 months eligible for this sample. We included 72% (13,061) of the sample regarding the explained variance for pain and hand function and excluded 28% (4973).

We used a fourth sample to assess the test-retest reliability. In January 2022, we invited 300 patients who had completed the mental health screening tool before clinician consultation to complete it again within 5 to 10 days. This had to be before their scheduled hand surgeon consultation. We included 35% (105) of the test-retest reliability sample. The median (range) time interval between measures was 6 days (5 to 10).

We assessed whether responders and nonresponders in the sample systematically differed regarding the association between the screening tool items and pain and hand function and the test-retest reliability. In the sample of the explained variance for pain and hand function, we defined responders as patients who completed all measures at baseline (sociodemographics and mental health questionnaires) and 3 months of follow-up (the VAS), whereas nonresponders were patients who only completed baseline measures. In the test-retest reliability sample, responders were patients who completed the primary test and retest, whereas nonresponders were patients who only completed the primary test. We calculated the standardized mean difference between responders and nonresponders. We only found small, clinically irrelevant differences in age and assigned treatment track between responders and nonresponders in the sample of the explained variance for pain and hand function (Supplemental Table 1 supplemental materials are available with the online version of CORR®). We found no differences between responders and nonresponders in the test-retest reliability sample (Supplemental Table 2 supplemental materials are available with the online version of CORR®)

Table 1. Baseline characteristics of the study samples

Variable	Sample 1: Test set (n = 18,034)	Sample 2: Validation set (n = 1017)	Sample 3: Association with pain and hand function (n = 13,061)	Sample 4: Test-retest reliability (n = 105)
Age in years	54 ± 15	57 ± 15	55 ± 14	56 ± 16
Sex = Female	65% (11,797)	63% (644)	66% (8602)	60% (63)
Treatment track				
Thumb regular	16% (2912)	-	15% (1977)	-
Thumb extended	7% (1191)	-	7% (964)	-
Dupuytren	9% (1561)	-	9% (1232)	-
Wrist regular	22% (3958)	-	20% (2669)	-
Wrist extended	8% (1388)	-	8% (1103)	-
Finger regular	19% (3441)	-	19% (2518)	-
Finger extended	3% (524)	-	3% (378)	-
Nerve (de-)compression	17% (3059)	-	17% (2220)	-
Duration of symptoms in months	19 ± 38	16 ± 30	19 ± 39	23 ± 66
Type of work				
Unemployed (including, retired)	34% (6110)	39% (396)	35% (4622)	42% (44)
Light physical labor (office work)	29% (5140)	24% (244)	28% (3699)	19% (20)
Moderate physical labor (working in a store)	27% (4790)	21% (217)	26% (3371)	29% (30)
Heavy physical labor (working in construction)	11% (1994)	16% (160)	11% (1369)	11% (11)
Treated/affected side[a]	41% (7455)	30% (308)	42% (5483)	33% (35)
Left	54% (9661)	37% (376)	54% (6995)	37% (39)
Right	5% (918)	33% (333)	5% (583)	30% (31)
Both				
Dominant hand	8% (1492)	10% (103)	8% (1076)	11% (12)
Left	89% (16039)	83% (841)	89% (11,061)	81% (85)
Right	3% (503)	7% (73)	3% (384)	8% (8)
Both				
Second opinion = No	96% (17,230)	85% (862)	95% (12,461)	89% (93)
PHQ-4 total score (scores can range from 0-12)	1.4 ± 2.3	1.7 ± 2.6	1.3 ± 2.2	-
PCS total score (scores can range from 0-52)	11.2 ± 9.7	13.5 ± 10.3	11 ± 9.5	-
B-IPQ total score (scores can range from 0-80)	37.0 ± 11.5	40.3 ± 11.0	36.8 ± 11.5	-

Data presented as % (n) or mean ± SD.

[a]For the validation set (sample 2) and the test-retest reliability sample (sample 4), the patient is asked which side is affected, whereas the values in sample 1 and 3 reflect the side that is treated.

After applying the eligibility criteria, we included 18,034 patients in the test set, 1017 patients in the validation set, 13,061 patients in the sample regarding the added value of the screening tool (that is, its association with pain and function), and 105 patients in the test-retest reliability sample (Fig. 1). The demographic characteristics of these patients were representative of a general population of patients with hand and wrist conditions (Table 1).

Variables, Data Sources, and Measurement

We measured pain catastrophizing using the 13-item PCS (score range 0 to 52; higher scores indicate more catastrophizing) [23], psychologic distress using the four-item PHQ-4 (score range 0 to 12; high scores indicate a potential anxiety or depression disorder) [24], and illness perception using the eight-item B-IPQ (total score range 0 to 80; higher scores indicate more negative illness perception) [25-27]. These are all valid and reliable instruments [23-27].

Sociodemographic characteristics collected at baseline included age, sex, measurement track (a predefined set of measurements at predefined timepoints based on the patient's diagnosis) [30], duration of symptoms, type of work, affected side, dominant hand, and whether a second opinion was sought. Lastly, we used the VAS, which is valid and reliable [31], to measure pain (range: 0 to 100, higher scores indicate more pain) and hand function (range: 0 to 100, lower scores indicate worse hand function) at baseline and 3 months.

Sample Size

Although large sample sizes (ideally more than 1000) [32] are required for chi-square automated interaction detection, we found no recommendations for sample size. Therefore, we used a convenience sample for the test set (postconsultation) and aimed to include more than 1000 participants. For the test-retest reliability sample, at least 50 participants are recommended [33], which is well below our sample of 105 participants.

Ethical Approval

Ethical approval for this study was obtained from Erasmus MC, Rotterdam, the Netherlands.

Statistical Analysis

We used a chi-square automated interaction detection [34] machine learning algorithm in the test set (Sample 1) to select the items for the screening tool. In each questionnaire, the chi-square automated interaction detection algorithm determined which item has the most discriminative power for the sum score of that questionnaire. These items were subsequently picked, and we calculated the Pearson correlation between these items and the associated sum score to assess the construct validity. To ensure high construct validity of the screening tool, we proposed that there should at least be a very strong correlation (that is, Pearson ≥ 0.80) [35] between the selected items and the sum score of the particular questionnaire in the test set for each construct of interest (such as pain catastrophizing, psychologic distress, or illness perception). We also calculated the Pearson correlation between the selected items and the sum scores in the validation set (Sample 2) to investigate whether the selected

items were solid, also aiming for a very strong correlation (that is, Pearson ≥ 0.80) in the validation set for the screening tool to be accurate.

To further assess the construct validity, we built linear regression models using Sample 3 to assess the explained variance of the screening tool items, with VAS pain during physical load and VAS hand function as dependent variables, both of which were measured at baseline and 3 months, adding up to four models. We built four additional models for the same dependent variables but with the total scores of the full questionnaires (the full PCS, PHQ-4, and B-IPQ) and compared the multiple R-squared of these models with those of the models only using the screening tool. In the models using the 3-month measurement as the dependent variable, we adjusted for baseline scores by adding the baseline score first in the model, because these are usually associated with the follow-up score [30]. By doing this, the explained variance we report is more reliably independent from the baseline scores.

For the test-retest reliability, we calculated the weighted kappa and ICCs for categorical items and ICCs and the standard error of measurement for continuous items.

Results

Screening Tool Development (Most Discriminative Item Selection and Correlation With Total Scores)

The chi-square automated interaction detection algorithm selected four items for the final screening tool (Table 2). For pain catastrophizing, the chi-square automated interaction detection algorithm selected item 4 (“When I’m in pain, it’s awful and I feel that it overwhelms me”) of the PCS (test set: Pearson correlation 0.82 [95% CI 0.81 to 0.82; p < 0.001], validation set 0.81 [0.79 to 0.83; p < 0.001], Fig. 2A). Item 2 of the PHQ-4 (“Not being able to stop or control worrying”) was selected for psychologic distress (test set 0.87 [95% CI 0.86 to 0.88; p < 0.001], validation set 0.88 [95% CI 0.86 to 0.89; p < 0.001], Fig. 2B). Two items of the B-IPQ were required to obtain a correlation greater than 0.80, resulting in the selection of items 6 (Concern: “How concerned are you about your illness?”) and 8 (Emotional response: How much does your illness affect you emotionally? (such as, does it make you angry, scared, upset or depressed?) (test set 0.85 [95% CI 0.85 to 0.86; p < 0.001], validation set 0.85 [95% CI 0.83 to 0.86; p < 0.001], Fig. 2C) from the chi-square automated interaction detection algorithm.

Table 2. The final screening tool for mental health

Item	Question	Score range	Response options
PCS item 4	When I'm in pain, it's awful and I feel that it overwhelms me	0-4	Not at all To a slight degree To a moderate degree To a great degree All the time
PHQ-4 item 2	Over the last 2 weeks, how often were you not able to stop or control worrying?	0-3	Not at all Several days More than half the days Nearly every day
B-IPQ item 6	How concerned are you about your illness?	0-10	Anchors: "Not at all concerned" (0) to "extremely concerned" (10)
B-IPQ item 8	How much does your illness affect you emotionally? (e.g., does it make you angry, scared, upset or depressed?	0-10	Anchors: "Not at all affected emotionally" (0) to "extremely affected emotionally" (10)

Construct Validity and Added Clinical Value (Association With Pain and Function)

The screening tool explained 17% of the variance in pain at baseline and 14% at 3 months. For function, this was 10% at baseline and 9% at 3 months. The full questionnaires performed only slightly better and explained 22% of the variance in pain at baseline and 15% at 3 months. For function, this was 13% at baseline and 10% at 3 months. Combined with the abovementioned correlations, this indicates the screening tool has high construct validity and added clinical value.

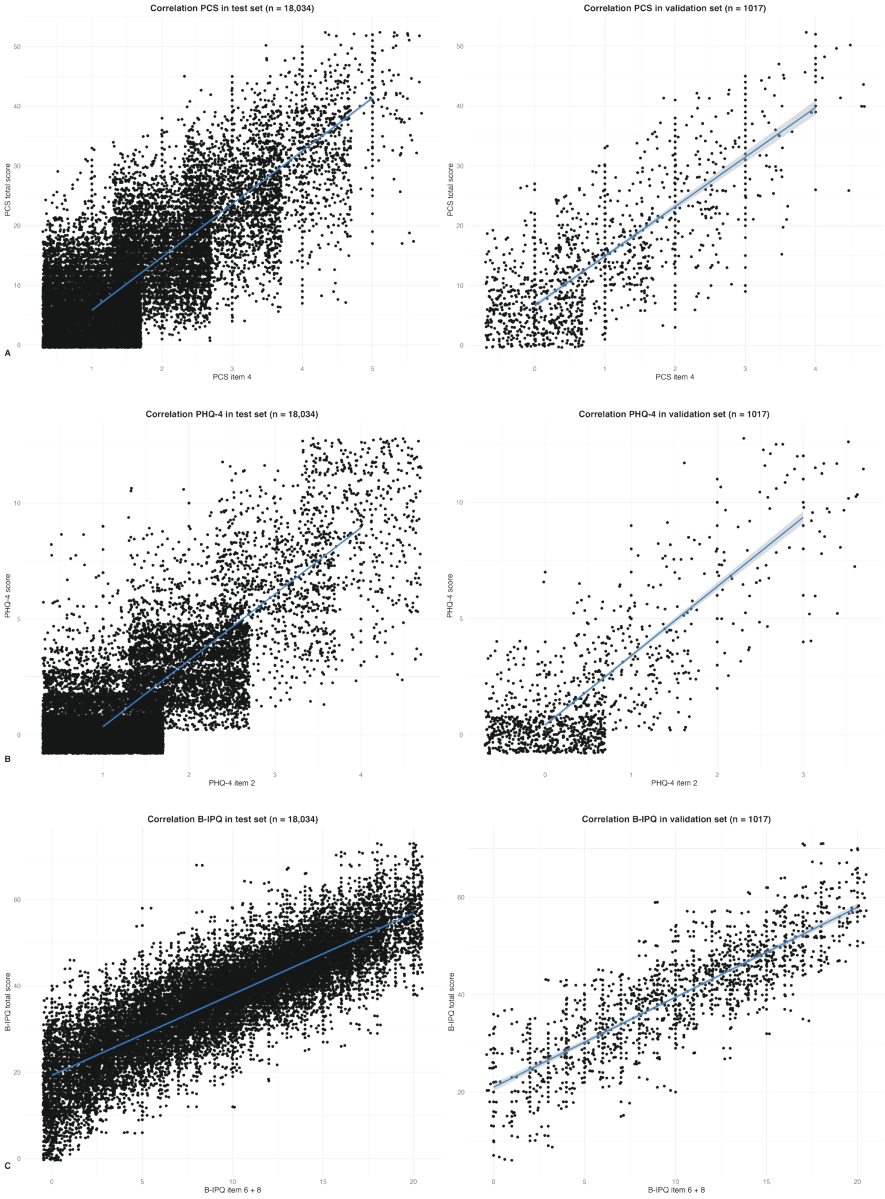


Fig. 2 These scatterplots demonstrate the correlation between the screening tool items and the sum scores. (A) For item 4 of the PCS and PCS total score, the Pearson correlation was 0.82 (95% CI 0.81 to 0.82; $p < 0.001$) in the test set (left) and 0.81 (95% CI 0.79 to 0.83; $p < 0.001$) in the validation set (right). (B) For item 2 of the PHQ-4 and PHQ-4 total score, the Pearson correlation was 0.87 (95% CI 0.86 to 0.88; $p < 0.001$) in the test set (left) and 0.88 (95% CI 0.86 to 0.89; $p < 0.001$) in the validation set (right). (C) For the B-IPQ items 6 and 8 and B-IPQ total score, the Pearson correlation was 0.85 (95% CI 0.85 to 0.86; $p < 0.001$) in the test set (left) and 0.85 (95% CI 0.83 to 0.86; $p < 0.001$) in the validation set (right).

Test-retest Reliability

There was a high test-retest reliability for PCS item 4 (ICC 0.75 [95% CI 0.66 to 0.83], weighted kappa 0.75 [95% CI 0.66 to 0.84]), B-IPQ items 6 (ICC 0.70 [95% CI 0.59 to 0.79]; standard error of measurement 1.4) and 8 (ICC 0.75 [95% CI 0.65 to 0.82]; standard error of measurement 1.3), whereas it was moderate for PHQ-4 item 2 (ICC 0.54 [95% CI 0.40 to 0.66], weighted kappa 0.54 [95% CI 0.38 to 0.70]) (Fig. 3A-D).

Response Time

The median total response time of the full PHQ-4, PCS, and B-IPQ was 4 minutes, 6 seconds. When assuming that the response time per item was equal among the questionnaires, the response time per item was 9 seconds for the PHQ-4, 8 seconds for the PCS, and 13 seconds for the B-IPQ. Given these assumptions, our newly developed screening tool has a response time of 43 seconds.

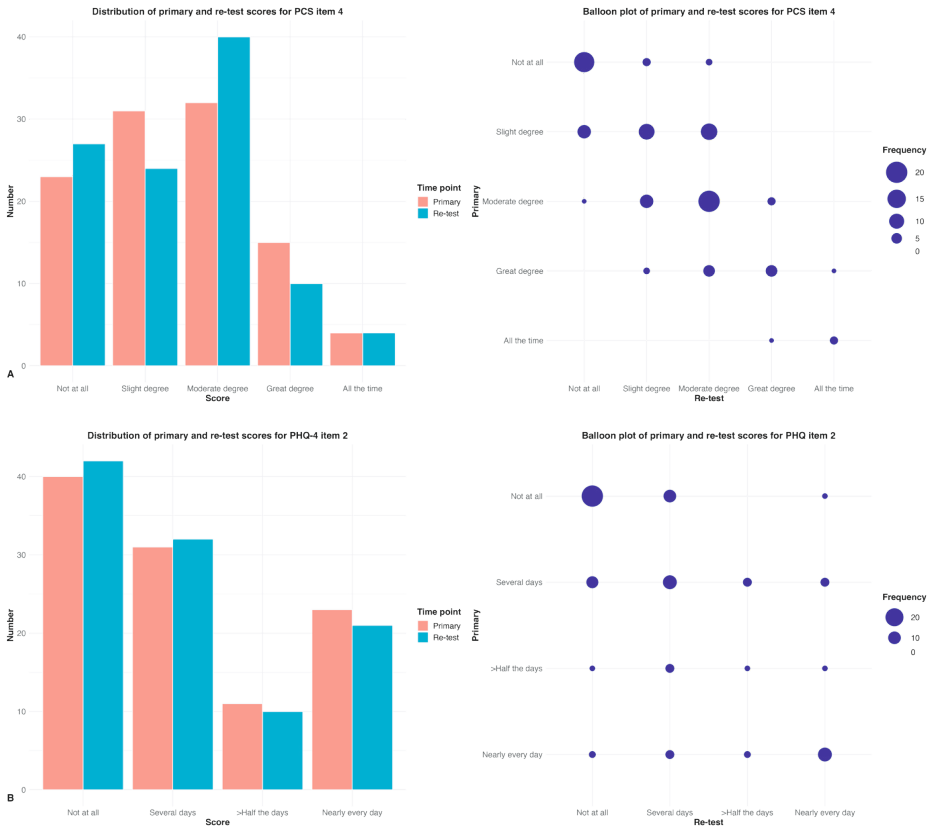


Fig. 3 These figures demonstrate the test-retest reliability of the screening tool items, which include (A) item 4 of the PCS, (B) item 2 of the PHQ-4, and (C-D) B-IPQ items 6 and 8. The bar plots demonstrate the score distribution in the primary test and the retest (left plots), whereas the balloon plots and the Bland-Altman plots demonstrate the discrepancy between the primary test and the retest of the screening tool items (right plots).

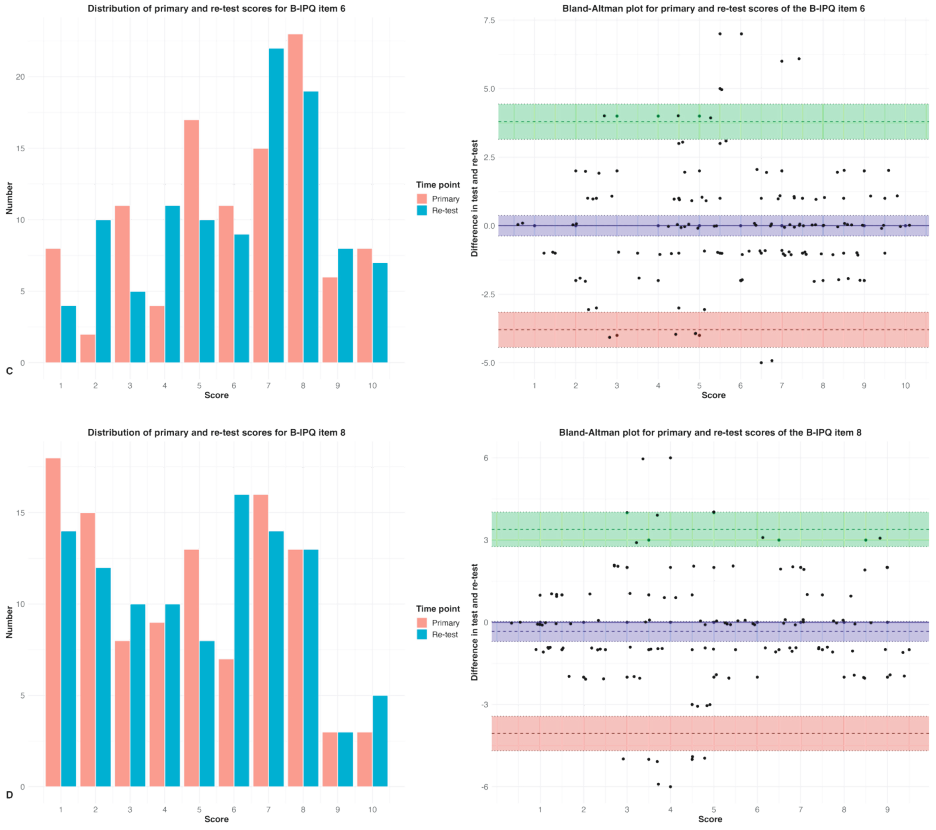


Fig. 3 *(continued)*

Discussion

Mental health has gained attention in musculoskeletal healthcare because it influences symptoms, outcomes, and decision-making. Measuring mental health in these patients can be challenging. There is a need for decision support tools that evaluate mental health in musculoskeletal healthcare. We developed a reliable and valid screening tool for pain catastrophizing, psychologic distress, and illness perception that contains only four questions and has an average response time of only 43 seconds. This tool can be used before clinician consultation and may serve as a conversation starter, an expectation management tool, or a decision support tool. For example, it may indicate that further testing is warranted, help guide a clinician when considering referral to a mental health specialist, or support a clinician in choosing between more invasive and less invasive treatments.

Limitations

A limitation inherent to our observational setting is missing data. However, because our nonresponder analysis suggested there were no clinically relevant differences between responders and nonresponders, we are confident this did not influence our results. Moreover, an advantage of our observational setting is the high ecological validity because our data reflect true daily practice.

Although there are screening tools for specific mental health constructs, there are, to the best of our knowledge, no other screening tools that aim for a combined measure of psychologic distress, pain catastrophizing, and illness perception. A limitation of our method is that we used two measurement instruments that have already been abbreviated (the four-item PHQ-4 for psychologic distress and the B-IPQ for illness perception) to select items for the screening tool. Therefore, our tool should only be used as an indication of one's mental health, and it should not be considered an in-depth mental health evaluation. However, the high construct validity of our screening tool indicates its items provide a valid view of the constructs of interest. Further, our tool was developed in patients with hand and wrist conditions, and although it seems generalizable, future research might investigate whether the tool can be used in different populations.

The screening tool has no normative values or cutoff scores. Although normative values or cutoff scores can be helpful in clinical decision-making, one may doubt if using these would be appropriate within our screening tool containing only a few questions. The constructs of interest are complex, and the aim of our tool is not to label patients in a certain category. Still, the scores of a patient on our screening tool provide much information that is helpful during clinical consultations, which may provide much context around the patient's symptoms.

Another limitation is that the estimated response time of the screening tool is calculated, not measured, and based on the assumption that the response time per item was equal in the full PCS, PHQ-4, and B-IPQ.

Discussion of Key Findings

The screening tool could serve as a conversation starter because it may reinforce a clinician's gut feeling about certain patients and could enable the clinician to discuss the patient's thoughts and feelings. Hypothetically, this may result in improved patient-reported experiences; for example, a patient may experience more clinician empathy if the clinician is attentive to the patient's thoughts and feelings. Additionally, it allows the clinician to manage expectations because these thoughts and feelings may affect treatment outcomes.

Because of the above, the screening tool may also serve as a decision support tool, because discussing these issues indicates that other treatment choices could be better. For example, if a patient with thumb-base osteoarthritis presents with high pain levels and

the mindset screening tool indicates a high degree of pain catastrophizing, a high degree of psychologic distress, and distorted illness perception, this indicates that possibly temporary decreased mental health may explain at least part of the patient's symptoms. In such case, a purely biomedical approach such as a thumb-base surgery may not be optimal, and less-invasive options may be considered first. Additionally, for example in this case, the screening tool might indicate whether an intervention into mental health may be beneficial, either performed by the clinician or a mental health specialist in more challenging cases. The above will only work if the screening tool is implemented and used, preferably before clinician consultation. Thus, future research could focus on implementing user-friendly data feedback to clinicians, such as through electronic dashboards. Additionally, studies might investigate whether using the tool yields other treatment choices (for example, changes in the ratio of invasive versus noninvasive treatment or the number of referrals to a mental health specialist), differences in outcome expectations, or differences in patient-reported experience measures. In line with this, future studies could also determine whether using the tool leads to better treatment outcomes, such as higher satisfaction with treatment results or increased cost-effectiveness.

The mental health screening tool explained a substantial part of the variance in pain and hand function at baseline and 3 months. This highlights, in line with other studies [1-18], the importance of mental health and its relation to treatment decisions and outcomes. The models with only the screening tool items performed nearly as well as the models using the full mental health measures (that is, the entire PCS, PHQ-4, and B-IPQ), which further substantiates the validity of our tool. Using the tool can reduce the time and patient burden of using patient-reported measures yet still collect relevant information for patient care and research.

Our screening tool had high test-retest reliability for most items. Only the PHQ-4 item yielded moderate test-retest reliability. Some other studies investigated the test-retest reliability of the PHQ-4 [36,37] and found better test-retest reliability than we did. However, these studies reported ICCs for the total score of the PHQ-4, whereas we assessed the test-retest reliability specifically for item 2 of the PHQ-4. It seems logical that the test-retest assessment of a single question yields more variability than a total score, because changes at an item level may cancel out at a total score level. Moreover, in our study, the test-retest reliability of this PHQ-4 item may also be affected by the fact that the item specifically asks for the degree of worrying in the past 2 weeks. Thus, hypothetically, the time interval between the test and the retest may also have caused an actual change in that item.

This four-item screening tool has a minimal time burden. If patient-reported measures of mental health are used at all in current daily practice, they are usually only distributed after clinician consultation. Especially with a screening tool that is this short, this is a missed opportunity, because treatment decisions are usually already made in this phase. Our data

indicate the screening tool can be reliably used before clinician consultation, which allows the screening tool we developed to be used in daily practice during clinician consultations.

Conclusion

This ultrashort, valid, and reliable screening tool for mental health (such as psychologic distress, pain catastrophizing, and illness perception) demonstrated added clinical value. The screening tool can be used in daily musculoskeletal healthcare practice as a conversation starter, an expectation management tool, or a decision support tool. For example, the screening tool may indicate that further testing is warranted, guide a clinician in referring to a mental health specialist, or support choices between more invasive and less invasive treatments. Future research could investigate in an experimental setting how this tool can be optimally used and whether using the tool yields other treatment choices or better outcome expectations, patient-reported experience measures, and treatment outcomes.

Group Authors

Members of the Hand-Wrist Study Group include: RAM Blomme, BJR Sluijter, DJJC van der Avoort, GJ Halbesma, niA Kroeze, J Smit, J Debeij, ET Walbeehm, GM van Couwelaar, GM Vermeulen, JP de Schipper, JFM Temming, JH van Uchelen, HL de Boer, KP de Haas, K Harmsen, OT Zöphel, R Feitz, JS Souer, R Koch, SER Hovius, TM Moojen, X Smit, R Hagen, R van Huis, PY Pennehouat, K Schoneveld, YE van Kooij, RM Wouters, J Veltkamp, A Fink, L Esteban Lopez, WA de Ridder, HP Slijper, RW Selles, JT Porsius, J Tsehaie, R Poelstra, MC Jansen, MJW van der Oest, L Hoogendam, JS Teunissen, JE Koopman, J Dekker, MHP ter Stege, JM Zuidam, CA van Nieuwenhoven, CA Hundepool, BEPA van der Heijden, JW Colaris, WR Bijlsma.

References

1. Barry MJ, Edgman-Levitan S. Shared decision making — the pinnacle of patient-centered care. *New Engl J Med.* 2012;366:780-781.
2. Basch E. Patient-reported outcomes — harnessing patients' voices to improve clinical care. *New Engl J Med.* 2017;376:105-108.
3. Becker SJ, Makarawung DJ, Spit SA, King JD, Ring D. Disability in patients with trapeziometacarpal joint arthrosis: Incidental versus presenting diagnosis. *J. Hand Surg. Am.* 2014;39:2009-2015.e2008.
4. Blackburn J, van der Oest MJW, Chen NC, et al. Are patient expectations and illness perception associated with patient-reported outcomes from surgical decompression in de quervain's tenosynovitis? *Clin Orthop Relat Res.* 2021;479:1147-1155.
5. Blackburn J, van der Oest MJW, Selles RW, et al. Which psychological variables are associated with pain and function before surgery for de quervain's tenosynovitis? A cross-sectional study. *Clin Orthop Relat Res.* 2019;477:2750-2758.
6. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res.* 2006;60:631-637.
7. Christodoulaki A, Baralou V, Konstantakopoulos G, Touloumi G. Validation of the patient health questionnaire-4 (phq-4) to screen for depression and anxiety in the greek general population. *J Psychosom Res.* 2022;160:110970.
8. Damman W, Liu R, Kaptein AA, et al. Illness perceptions and their association with 2 year functional status and change in patients with hand osteoarthritis. *Rheumatology (Oxford).* 2018;Dec 1;57(12):2190-2199. doi: 10.1093/rheumatology/key231.
9. Dave AJ, Selzer F, Losina E, et al. Is there an association between whole-body pain with osteoarthritis-related knee pain, pain catastrophizing, and mental health? *Clin. Orthop. Relat. Res.* 2015;473:3894-3902.
10. de Raaij EJ, Schröder C, Maissan FJ, Pool JJ, Wittink H. Cross-cultural adaptation and measurement properties of the brief illness perception questionnaire-dutch language version. *Man Ther.* 2012;17:330-335.
11. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine: A practical guide.* Cambridge University Press; 2011.
12. Duivenvoorden T, Vissers MM, Verhaar JAN, et al. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: A prospective multicentre study. *Osteoarthritis Cartilage.* 2013;21:1834-1840.
13. Evans JD. *Straightforward statistics for the behavioral sciences.* Thomson Brooks/Cole Publishing Co; 1996.
14. Foster NE, Thomas E, Hill JC, Hay EM. The relationship between patient and practitioner expectations and preferences and clinical outcomes in a trial of exercise and acupuncture for knee osteoarthritis. *Eur. J. Pain.* 2010;14:402-409.
15. Frouzakis R, Herren DB, Marks M. Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis. *J. Hand Surg. Am.* 2015;40:483-490.

16. Glette M, Landmark T, Jensen MP, et al. Catastrophizing, solicitous responses from significant others, and function in individuals with neuropathic pain, osteoarthritis, or spinal pain in the general population. *J. Pain.* 2018;19:983-995.

17. Hawker GMSKTFM. Measures of adult pain: Visual analog scale for pain (vas pain), numeric rating scale for pain (nrs pain), mcgill pain questionnaire (mpq), short-form mcgill pain questionnaire (sf-mpq), chronic pain grade scale (cpgs), short form-36 bodily pain scale (sf-36 bps), and measure of intermittent and constant osteoarthritis pain (icoap). *Arthritis Care Res. (Hoboken).* 2011;63 Suppl 11:S240-S252.

18. Hoogendam L, van der Oest MJW, Tsehaie J, et al. Psychological factors are more strongly associated with pain than radiographic severity in non-invasively treated first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021;43:1897-1902.

19. Hoogendam L, van der Oest MJW, Wouters RM, et al. Patients with higher treatment outcome expectations are more satisfied with the results of nonoperative treatment for thumb base osteoarthritis: A cohort study. *Arch Phys Med Rehabil.* 2021;102:1533-1540.

20. Kass GV. An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society. Series C (Applied Statistics).* 1980;29:119-127.

21. Kim HW, Shin C, Lee SH, Han C. Standardization of the korean version of the patient health questionnaire-4 (phq-4). *Clin Psychopharmacol Neurosci.* 2021;19:104-111.

22. Kroenke K, Spitzer Rl - Williams JBW, Williams Jb - Lowe B, Lowe B. An ultra-brief screening scale for anxiety and depression: The phq-4. *Psychosomatics.* 2009;50:613–621.

23. Leysen M, Nijs J, Meeus M, et al. Clinimetric properties of illness perception questionnaire revised (ipq-r) and brief illness perception questionnaire (brief ipq) in patients with musculoskeletal disorders: A systematic review. *Man Ther.* 2015;20:10-17.

24. Porter ME. What is value in health care? *N Engl J Med.* 2010;363:2477-2481.

25. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. *Clin Orthop Relat Res.* 2010;468:798-806.

26. Şata M, Elkonca F. A comparison of classification performances between the methods of logistics regression and chaid analysis in accordance with sample size. *Int. J. Contemp. Educ. Res.* 2020;7:15-26.

27. Selles RW, Wouters RM, Poelstra R, et al. Routine health outcome measurement: Development, design, and implementation of the hand and wrist cohort. *Plast Reconstr Surg.* 2020.

28. Sharma A, Kudesia P, Shi Q, Gandhi R. Anxiety and depression in patients with osteoarthritis: Impact and management challenges. *Open access rheumatology : research and reviews.* 2016;8:103-113.

29. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: Development and validation. *Psychol. Assess.* 1995;7:524-532.

30. van der Oest MJW, Hoogendam L, Wouters RM, et al. Associations between positive treatment outcome expectations, illness understanding, and outcomes: A cohort study on non-operative treatment of first carpometacarpal osteoarthritis. *Disabil Rehabil.* 2021:1-8.

31. van der Oest MJW, Poelstra R, Feitz R, et al. Illness perceptions of patients with first carpometacarpal osteoarthritis, carpal tunnel syndrome, dupuytren contracture, or trigger finger. *J Hand Surg Am.* 2020;45:455.e451-455.e458.

32. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (strobe) statement: Guidelines for reporting observational studies. *Lancet.* 2007;370:1453-1457.

33. Wood TJ, Thornley P, Petruccelli D, Kabali C, Winemaker M, de Beer J. Preoperative predictors of pain catastrophizing, anxiety, and depression in patients undergoing total joint arthroplasty. *J. Arthroplasty.* 2016;31:2750-2756.

34. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, Hovius SER. A standard set foroutcome measurement in patients with hand and wrist conditions: Consensus by the international consortium for health outcomes measurement hand and wrist working group. *J Hand Surg Am.* 2021;46:841-855.e847.

35. Wouters RM, Porsius JT, van der Oest MJW, et al. Psychological characteristics, female sex, and opioid use predict acute postoperative pain in patients surgically treated for thumb base osteoarthritis: A cohort study. *Plast Reconstr Surg.* 2020;146:1307-1316.

36. Wouters RM, Vranceanu AM, Slijper HP, et al. Patients with thumb-base osteoarthritis scheduled for surgery have more symptoms, worse psychological profile, and higher expectations than nonsurgical counterparts: A large cohort analysis. *Clin Orthop Relat Res.* 2019;477:2735-2746.

37. Wouters RM, Vranceanu AM, Slijper HP, et al. Patients with thumb-base osteoarthritis scheduled for surgery have more symptoms, worse psychological profile, and higher expectations than nonsurgical counterparts: A large cohort analysis. *Clin Orthop Relat Res.* 2019.

Supplementary Table 1. Nonresponder analysis of Sample 3 (the association with pain and hand function)

Variable	Responders (n = 13,061)	Nonresponders (n = 4,973)	Standardized mean difference
Age	55 ± 14	51 ± 16	0.247
Sex = Female	66% (8602)	64% (3195)	0.034
Treatment track			
Thumb regular	15% (1977)	19% (935)	0.236
Thumb extended	7% (964)	5% (227)	
Dupuytren	9% (1232)	7% (329)	
Wrist regular	20% (2669)	26% (1289)	
Wrist extended	8% (1103)	6% (285)	
Finger regular	19% (2518)	19% (923)	
Finger extended	3% (378)	3% (146)	
Nerve (de-)compression	17% (2220)	17% (839)	
Duration of symptoms in months	19 ± 39	18 (35)	0.031
Type of work			
Unemployed (retired)	35% (4622)	30% (1488)	0.126
Light physical labor (office work)	28% (3699)	29% (1441)	
Moderate physical labor	26% (3371)	29% (1419)	
(working in a store)	11% (1369)	13% (625)	
Heavy physical labor (working in construction)			
Treated/affected side[a]			
Left	42% (5483)	40% (1972)	0.103
Right	54% (6995)	54% (2666)	
Both	5% (583)	7% (335)	
Dominant hand			
Left	8% (1076)	8% (416)	0.034
Right	89% (11,061)	89% (4438)	
Both	3% (384)	2% (119)	
Second opinion = No	95% (12,461)	96% (4769)	0.024
PHQ-4 total score (scores can range from 0-12)	1.3 ± 2.2	1.5 ± 2.4	0.073
PCS total score (scores can range from 0-52)	11 ± 9.5	11.7 ± 10.2	0.065
B-IPQ total score (scores can range from 0-80)	36.8 ± 11.5	37.4 ± 11.	0.047

Data presented as mean ± SD or % (n).

Responders are defined as patients who completed both the measures of interest at baseline (that is, sociodemographics and mental health questionnaires) and 3 months follow-up (the VAS), whereas nonresponders are defined as patients that only completed the measures at baseline. [a]For the validation set (sample 2) and the test-retest reliability sample (sample 4), the patient is asked which side is affected, whereas the values in sample 1 and 3 reflect the side that is treated.

Supplementary Table 2. Nonresponder analysis of Sample 4 (test-retest reliability)

Variable	Responders (n = 105)	Nonresponders (n = 195)	Standardized mean difference
Age in years	56 ± 16	55 ± 16	0.073
Sex is Female	60% (63)	61% (119)	0.021
Duration of symptoms in months	23 ± 66	14 ± 20	0.194
Type of work			
Unemployed (retired)	42% (44)	37% (72)	0.192
Light physical labor (office work)	19% (20)	26% (50)	
Moderate physical labor (working in a store)	29% (30)	25% (48)	
Heavy physical labor (working in construction)	11% (11)	13% (25)	
Treated/affected side[a]	33% (35)	30% (59)	0.105
Left	37% (39)	42% (82)	
Right	30% (31)	27% (53)	
Both			
Dominant hand	11% (12)	10% (19)	0.065
Left	81% (85)	82% (159)	
Right	8% (8)	9% (17)	
Both			
Second opinion = No	89% (93)	89% (173)	0.019

Data presented as mean ± SD or % (n).

Responders are defined as patients that completed both the primary test and the retest, whereas nonresponders are defined as patients that only completed the primary test. [a]For the validation set (sample 2) and the test-retest reliability sample (sample 4), the patient is asked which side is affected, whereas the values in sample 1 and 3 reflect the side that is treated.


6

7

TAILORING AND EVALUATING TREATMENT WITH THE PATIENT-SPECIFIC NEEDS EVALUATION: A PATIENT-CENTERED APPROACH

de Ridder WA, van Kooij YE, Slijper HP, Arends GR, de Roode A, MacDermid JC, Vermeulen GM, Hovius SER, Selles RW, The Hand-Wrist Study Group, Wouters RM;
2023; Plastic and Reconstructive Surgery

Abstract

Background

No patient-reported instrument assesses patient-specific information needs, treatment goals, and Personal Meaningful Gain (PMG, a novel construct evaluating individualized, clinically relevant improvement). This study reports the development of the Patient-Specific Needs Evaluation (PSN) and examines its discriminative validity (i.e., its ability to distinguish satisfied from dissatisfied patients) and test-retest reliability in patients with hand or wrist conditions.

Methods

A mixed-methods approach was used to develop and validate the PSN, following COSMIN guidelines, including pilot testing, a survey (pilot: n=223, final PSN: n=275), cognitive debriefing (n=16), expert input, and validation. Discriminative validity was assessed by comparing the satisfaction level of patients who did or did not achieve their PMG (n=1,985) and test-retest reliability using absolute agreement, Cohen's kappa, and ICCs (n=102). We used a sample of 2,860 patients to describe responses to the final PSN.

Results

The PSN has only five questions (completion time ±3 minutes) and is freely accessible online. The items and response options were considered understandable by 90-92% and complete by 84-89% of the end-users. The PSN had excellent discriminative validity (Cramer's V: 0.48, $p<0.001$) and moderate to high test-retest reliability (Kappa: 0.46-0.68, ICCs: 0.53-0.73).

Conclusions

The PSN is a freely available patient-centered decision-support tool that helps clinicians tailor their consultations to the patient's individual needs and goals. It contains the PMG, a novel construct evaluating individualized, clinically relevant treatment outcomes. The PSN may function as a conversation starter, facilitate expectation management, and aid shared decision-making. The PSN is implementation-ready and can be readily adapted to other patient populations.

Introduction

Patient-centered and value-based healthcare frameworks have gained global recognition in recent years, aiming to put the patient first and achieve better outcomes at lower costs[1-4]. Key in these frameworks is responding to individual information needs and treatment goals[5], aiming for high satisfaction with the treatment results[6-12]. It is, therefore, important for clinicians to be well-informed about the patient's information needs and treatment goals. Clinicians aim to meet patients' needs and goals, but sometimes a misalignment occurs. For instance, a surgeon may prioritize pain relief with a joint replacement while the patient prioritizes hand appearance. This misalignment can induce a treatment plan not fully meeting the patient's needs or goals.

In routine care, clinicians must understand each patient's information needs, as patients require information to comprehend their medical situation, participate in decision-making, and manage expectations. Providing targeted, patient-specific information improves shared decision making[13], daily functioning[14], treatment adherence[15], quality of life, the patient's mindset, pretreatment expectations[16-24], and satisfaction with care and treatment results[25]. Since information provision is modifiable[26-29], outcomes can be improved. There is a lack of concise, patient-reported tools that focus on patients' information needs and treatment goals. These needs and goals may be, for example, understanding the diagnosis or regaining the ability to perform daily activities. Setting goals enhances patient participation, treatment adherence, and motivation, ultimately improving outcomes and satisfaction with treatment results[30-32]. There are several limitations to existing tools assessing patient-specific limitations or goals, including the Canadian Occupational Performance Measure[33], Goal Attainment Scaling[34], Patient-Specific Goalsetting Method[35], and the Patient-Specific Functional Scale[36]. These limitations depend on the specific tool and include being time-consuming[32], having the potential for therapist bias[33-35], and only focusing on the activities and participation levels instead of all levels of the International Classification of Functioning, Disability, and Health (ICF)[33-37]. Moreover, they do not assess patient-specific improvement goals, i.e., when is the patient satisfied with the treatment results? Patient-specific improvement goals may depend on condition, treatment type, baseline score, and other patient-specific factors. For example, a recreational tennis player may consider a change from 4 to 8 on a 0-10 scale satisfactory, whereas a professional tennis player may not. We introduce the Personal Meaningful Gain (PMG) to represent the improvement an individual wants to obtain in a domain relevant to that individual, given the baseline score. Knowing 1) the information needs, 2) the individual goal, and 3) the PMG before treatment will allow clinicians to improve decision support and facilitate expectation management.

This study introduces a brief patient-reported tool assessing patient-specific information needs, treatment goals, and Personal Meaningful Gain before a first clinician consultation: the Patient-Specific Needs Evaluation (PSN). Specifically, the first objective of this study

was to describe the development of the brief, easy-to-use, patient-reported tool to assess 1) patient-specific information needs, 2) treatment goals, and 3) Personal Meaningful Gain (PMG). This tool was initially developed for patients with hand and wrist conditions, but we designed it to be easily adopted in other patient populations. The second study objective was to examine the PSN's discriminative validity (i.e., its ability to distinguish satisfied from dissatisfied patients) and test-retest reliability. The third study objective was to describe the results of the final PSN.

Methods

Study design

This was a user-centered mixed-methods study in patients with hand or wrist conditions, healthcare providers, and other stakeholders. We used the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines on PROM development[38] and measurement properties[38].

Setting

We developed the Patient-Specific Needs Evaluation (PSN) at Erasmus MC (an academic hospital) and Xpert Clinics (a specialized clinic for hand and wrist care) in The Netherlands. Data were collected at Xpert Clinics[21] between July and August 2023. The medical ethical review committee of Erasmus MC approved this study; all participants provided informed consent.

Research team

The core research team consisted of hand surgeons and therapists (WR, YK, RW, SH, GRA, AR, GV, JMD), professionals with experience in developing measurement sets and tools (RW, SH, HS, JMD, RS)[12,39-42], and electronic data capturing and implementation (HS, YK, RW, RS, SH, JMD, GV, WR)[21,43]. We consulted other clinicians, language experts, and native English speakers.

PSN development process

The development was iterative and comprised five overlapping stages, each informing subsequent stage(s) (Figure 1). Stage 1 included literature studies and expert meetings. After developing an item bank, we conducted a pilot study and survey on completeness and understandability in Stage 2. Stage 3 included cognitive debriefing of patients and clinicians and refining the item bank. We gathered expert input in Stage 4 and consulted a language expert, performed cross-cultural translation, and repeated the survey for the final PSN in Stage 5 (more details in Figure 1).

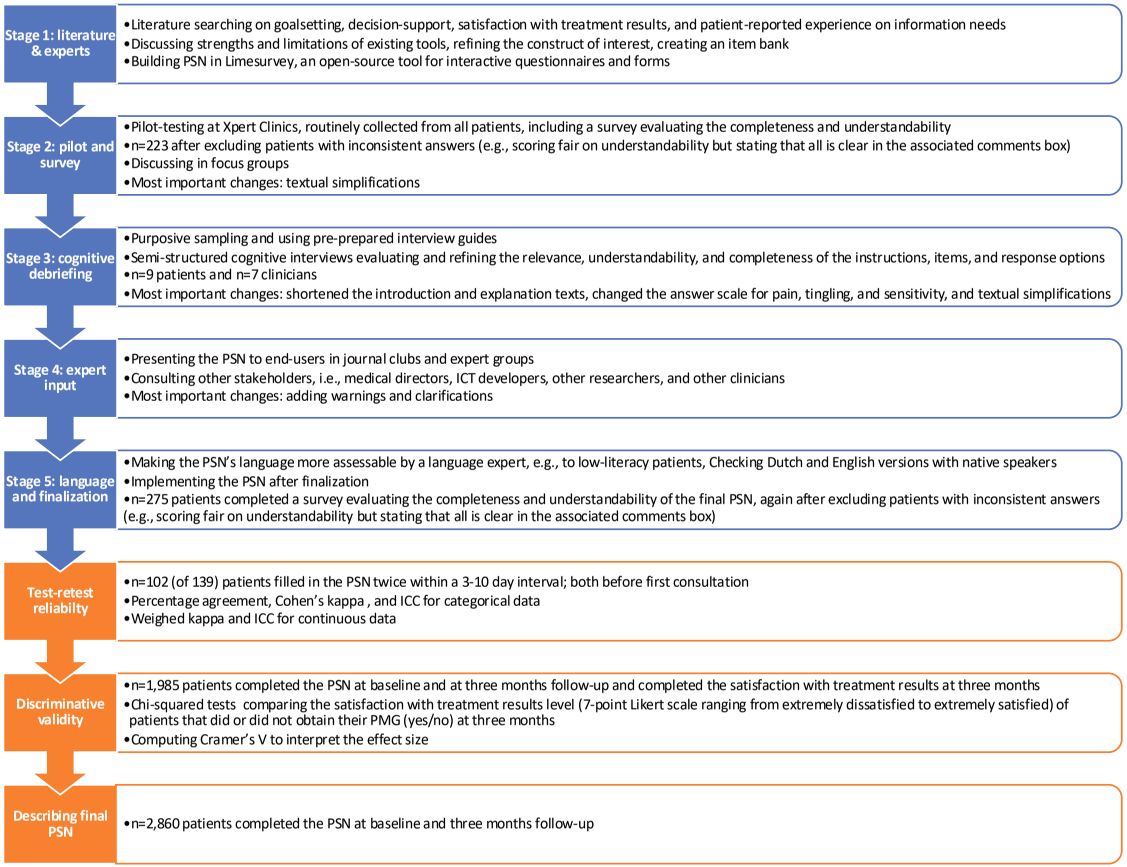


Fig 1. Flow chart of the development (in blue) and the validation (in orange) of the PSN, describing the sample and most important goals and activities per stage.

Participants

We used different samples to develop the PSN and establish the discriminative validity and test-retest reliability (Figure 1). For all samples, patients were eligible if they were adults, had any hand or wrist condition, completed our intake questionnaire, and understood Dutch language. All questionnaires were completed digitally.

For the survey, we excluded patients who gave inconsistent answers, e.g., stating fair on understandability but stating that all is clear in the associated comments box.

For discriminative validity, we included patients who completed the PSN at baseline and three months follow-up, as well as the Satisfaction with Treatment Results Questionnaire at three months[11,12]. We prospectively invited patients to participate in a test-retest study and complete the PSN for a second time 3-5 days after initial completion. The retest remained accessible for six days, i.e., a possible time interval of 3-11 days. We hypothesized that patient needs and goals remained stable during this period. We included patients in the test-retest analysis if they completed both the primary and retest PSN before clinician consultation. COSMIN advices a sample size of >100 participants for examining test-retest reliability[44]. To describe the results of the final PSN, we included all patients that completed the PSN at baseline and three months follow-up. There were no additional exclusion criteria. All samples reflected the target population (patients with hand and wrist conditions) and differed in age, sex, and treatment location.

Discriminative validity, test-retest reliability, and statistical analysis

We evaluated discriminative validity by comparing the satisfaction with treatment results level of patients that did or did not obtain their PMG. We used a Satisfaction with Treatment Results Questionnaire[11,12] at three months, which evaluates satisfaction using a 7-point Likert scale, ranging from extremely dissatisfied to extremely satisfied. Using Chi-squared tests, we determined the PMG's discriminative power. We computed Cramer's V to interpret the effect size, where 0.10 reflects a small effect size, 0.30 a medium effect size, and 0.50 a large effect size[45].

We evaluated test-retest reliability by computing absolute agreement and Cohen's kappa. We computed intraclass correlation coefficients (ICCs) for all variables, including the goal domain, baseline score, the score needed to be satisfied with the most important goal domain, and the PMG. Kappa scores lie between -1 and 1, where ≤ 0 indicates no agreement, 0.01-0.20 none to slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81-1.00 is almost-perfect agreement[46]. We calculated ICCs using a two-way mixed-effects model[47]. ICCs range from 0 to 1, 1 being perfect reliability, 0.90-0.99 very high, 0.70-0.89 high, 0.50-0.69 moderate, 0.26-0.49 low, and 0.00-0.25 indicates little if any, reliability[48-50].

There were no missing data in the final PSN, as completing it before clinician consultation is mandatory in our clinical setting. We analyzed missing data patterns for the test-retest

analyses; patients who completed both the primary and retest tests were responders, whereas patients without a retest were non-responders. We compared baseline characteristics of responders and non-responders using significance testing and calculating standardized mean differences to investigate if they systematically differed. We used R statistical software version 4.1.1 for the quantitative analyses and considered a p-value <0.05 significant. We tested the Dutch version of the PSN.

Results

Development process: cognitive debriefing and survey data

We performed sixteen cognitive interviews among nine patients and seven clinicians. All patients (three men, six women, aged 21-71 years (median: 51 years)) had different diagnoses. We also included patients with lower levels of education. Amongst clinicians, we interviewed one occupational hand therapist, two physical hand therapists, and four hand surgeons (five men, two women, aged 27-70 years (median: 40 years)). We iteratively improved the PSN, alternating between interviewing and adjusting, e.g., we shortened the introduction and explanation texts, changed the answer scale for pain, tingling, and sensitivity, and simplified the text with a language expert (Supplemental Digital Content (SDC) 1-2).

The survey on the final PSN indicated that the questions and response options were rated entirely or mostly understandable by 90-92% and fully or mostly complete by 84-89% of the 275 participants (SDC 3A-F). These were 89-93% and 86-91% for the pilot PSN (n=223).

The final PSN

Because of the dependencies within the PSN, it works best in digital form. It can be accessed here: https://personeel.equipezorgbedrijven.nl/ls/index.php?r=survey/index&sid=587344&lang=en (See Table 1 for a non-digital version). The intake PSN has five questions and takes approximately three minutes to complete. The information need part asks an open question about the patient's reason for making an appointment at the clinic (their request for help), followed by a single-select question where respondents pick their most important information need category. Subsequently, respondents select a predefined sub-answer based on that category to specify their information need in more detail. The treatment goal part of the PSN asks respondents to choose which domain they would most like to improve if they were to be treated and rate their baseline score on that domain on a 0-10 scale (e.g., the baseline pain score). Two secondary goal domains can optionally be selected. The final question asks for the score they think they need to achieve with treatment to be satisfied. The Personal Meaningful Gain (PMG) is then automatically generated as the difference between the respondent's baseline performance rating and their score needed to be satisfied (Figure 2). The follow-up PSN evaluates the previously selected information needs and treatment and improvement goals in only two questions and takes less than one minute to complete.

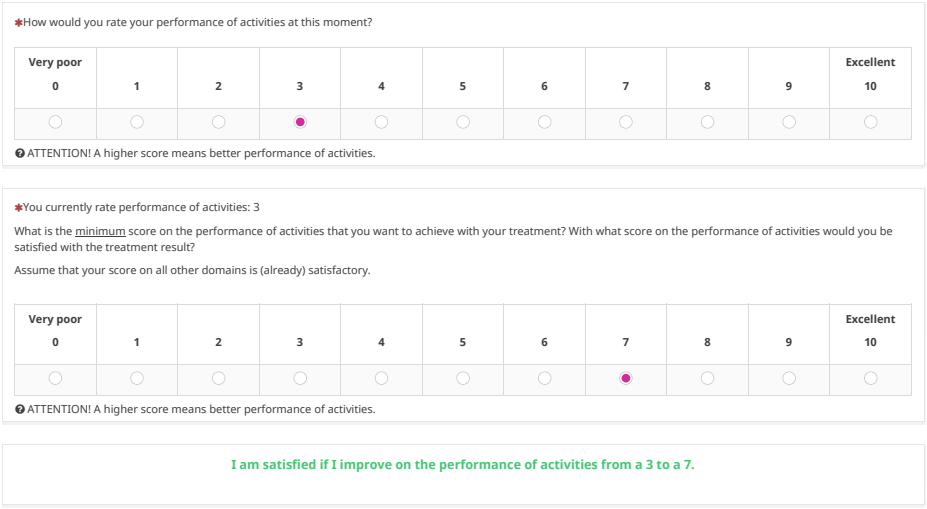


Fig. 2: Visualization of the Patient-Specific Needs (PSN) treatment goal and Personal Meaningful Gain (PMG) parts. In this example, the patient entered that the most important treatment goal was to improve the performance of activities. The score at baseline was 3 on a 0-10 scale (high scores indicate better performance), and the patient indicated that a score of 7 is needed to become satisfied with the treatment result. After this is filled in, the digital PSN automatically generates a statement on the treatment goal, and PMG, for the patient to be able to check whether it is correct or needs modification.

The final PSN was completed by 2,860 patients (Table 2). Figure 3 shows the selected information need categories, and Figure 4 shows the distribution of the selected treatment goals. The rating on the most important domain was normally distributed with a median of 4 (Figure 5). The median score needed to be satisfied with the treatment result was 9 (Figure 5).

Discriminative validity and test-retest reliability

We included 1985 patients for the discriminative validity analysis (Table 2). Patients who obtained the PMG had better satisfaction with treatment results than those who did not (Figure 6, $p<0.001$). There was a medium to large effect size (Cramer's V: 0.48), indicating that the PMG has excellent discriminative validity, i.e., the ability to distinguish satisfied from dissatisfied patients.

Table 1. The non-digital version of the Patient-Specific Needs questionnaire (PSN). The PSN is best administered in digital form and can be accessed digitally and open access here: https://personeel.equipezorgbedrijven.nl/ls/index.php?r=survey/index&sid=587344&lang=en. This table displays each question and the associated response options, which, in some specific domains, are slightly different than displayed. After question 4, respondents can optionally pick two secondary domains.

Part	Question	Response options
Information needs	1. What is the reason that you have made an appointment with us? In other words: what is your request for help from the doctor?	[Open text]
	2A. What is your most important information need?	Choose one of the following options: I do not need information Diagnosis (I have questions about the diagnosis) Advice (I want to know what is the best thing to do in my situation) Treatment (I have questions about the treatment) Perspective (I want to know what to expect in the future)
	2B. Specifying question based on information need: On which topic would you like advice? OR What would you like to know about the diagnosis? OR What would you like to know about the treatment? OR What would you like to know about your perspective?	[Choose one of the response options dependent on information need category, see digital PSN for all options]
Treatment and improvement goals	3. If you were treated, which domain would you most like to improve?	Choose one of the following options: I do not want to be treated Numbness (loss of sensation) Mobility / flexibility of my hand Strength Pain Tingling Performance of activities (e.g., housekeeping, hobby, sports...) Appearance of my hand / wrist Ability to work

Table 1. *(continued)*

Part	Question	Response options
	4. How would you rate your [domain from question 3] at this moment?	Score range 0-10; higher scores indicate better performance except for the items "Numbness (loss of sensation)", "Pain", and "Tingling"
	5. What is the minimum score on [domain] that you want to achieve with your treatment? With what score would you be satisfied with the treatment result? Assume that your score on all other domains is (already) satisfactory.	Score range 0-10; higher scores indicate better performance except for the items "Numbness (loss of sensation)", "Pain", and "Tingling"

For the test-retest reliability, 102 of the 139 invited patients completed both the primary test and the retest within a median interval of 7 days (range 3-11 days). We found small differences between responders and non-responders in age and type of work (SDC 4). There was moderate agreement and reliability for the most important goal domain (Table 3, SDC 5). Considering it also agreement when the most important goal domain was chosen as a secondary goal domain in the retest, the test-retest improved to substantial agreement and high reliability (Table 3, SDC 6). We found moderate reliability for the baseline score on the most important goal domain, for the score needed to be satisfied, and the PMG (Table 3).

Table 2. Baseline characteristics of the patients that completed the final PSN (n = 2,860), patients that participated in the test-retest sample (n = 102), and in the discriminative validity sample (n = 1,985).

Variable	Sample that completed the final PSN (n = 2,860)	Discriminative validity sample (n = 1,985)	Test-retest sample (n = 102)
Age, mean (SD)	54 (16.3)	59 (13.9)	61 (15.5)
Sex = male, n (%)	1086 (38.0)	704 (35.5)	46 (45.1)
Duration of symptoms in months, mean (SD)	18 (38.2)	17 (33.5)	21 (39.6)
Type of work, n (%)			
Unemployed due retirement	695 (24.3)	570 (28.7)	41 (40.2)
Unemployed due other reason	339 (11.9)	214 (10.8)	6 (5.9)
Light physical labor (e.g., office work)	735 (25.7)	468 (23.6)	22 (21.6)
Moderate physical labor (e.g., working in a store)	648 (22.7)	438 (22.1)	16 (15.7)
Heavy physical labor (e.g., working in construction	443 (15.5)	295 (14.9)	17 (16.7)

Table 2. *(continued)*

Variable	Sample that completed the final PSN (n = 2,860)	Discriminative validity sample (n = 1,985)	Test-retest sample (n = 102)
Level of education (%)			
None	34 (1.2)	12 (0.6)	1 (1.0)
Primary education (primary school, special primary education)	71 (2.5)	31 (1.6)	1 (1.0)
Primary or pre-vocational education (such as (in Dutch) LTS, LEAO, LHNO, Huishoudschool, VMBO)	323 (11.3)	252 (12.7)	12 (11.8)
Secondary general secondary education (such as (in Dutch) MAVO, (M)ULO, MBO-short, VMBO-t)	517 (18.1)	356 (17.9)	24 (23.5)
Secondary vocational education and vocational training (such as (in Dutch) MKBO-long, MTS, MEAO, BOL, BBL, INAS)	599 (20.9)	429 (21.6)	20 (19.6)
Higher general and pre-university education (such as (in Dutch) HAVO, VWO, Atheneum, Gymnasium, HBS, MMS)	251 (8.8)	198 (10.0)	9 (8.8)
Higher vocational education (such as (in Dutch) HBO, HTS, HEAO, HBO-V, university graduates	608 (21.3)	466 (23.5)	21 (20.6)
Scientific education (e.g., MSc.)	299 (10.5)	164 (8.3)	8 (7.8)
Prefer not to say	158 (5.5)	77 (3.9)	6 (5.9)
Body Mass Index, mean (SD)	26.5 (4.7)	27.2 (4.9)	26.5 (4.4)
Smoking status, n (%)			
Yes, daily smoker	367 (12.8)	207 (10.4)	10 (9.8)
Yes, passive smoker	15 (0.5)	8 (0.4)	2 (2.0)
Yes, sometimes	140 (4.9)	76 (3.8)	6 (5.9)
No	2338 (81.7)	1694 (85.3)	84 (82.4)
Affected side, n (%)			
Left	930 (32.5)	607 (30.6)	33 (32.4)
Right	1106 (38.7)	743 (37.4)	40 (39.2)
Both	824 (28.8)	635 (32.0)	29 (28.4)
Dominance, n (%)			
Left	299 (10.5)	199 (10.0)	11 (10.8)
Right	2395 (83.7)	1676 (84.4)	84 (82.4)
Both	166 (5.8)	110 (5.5)	7 (6.9)
Second opinion = no, n (%)	2475 (86.5)	1781 (89.7)	87 (85.3)
Personal injury lawsuit = no, n (%)	2801 (97.9)	1960 (98.7)	100 (98.0)

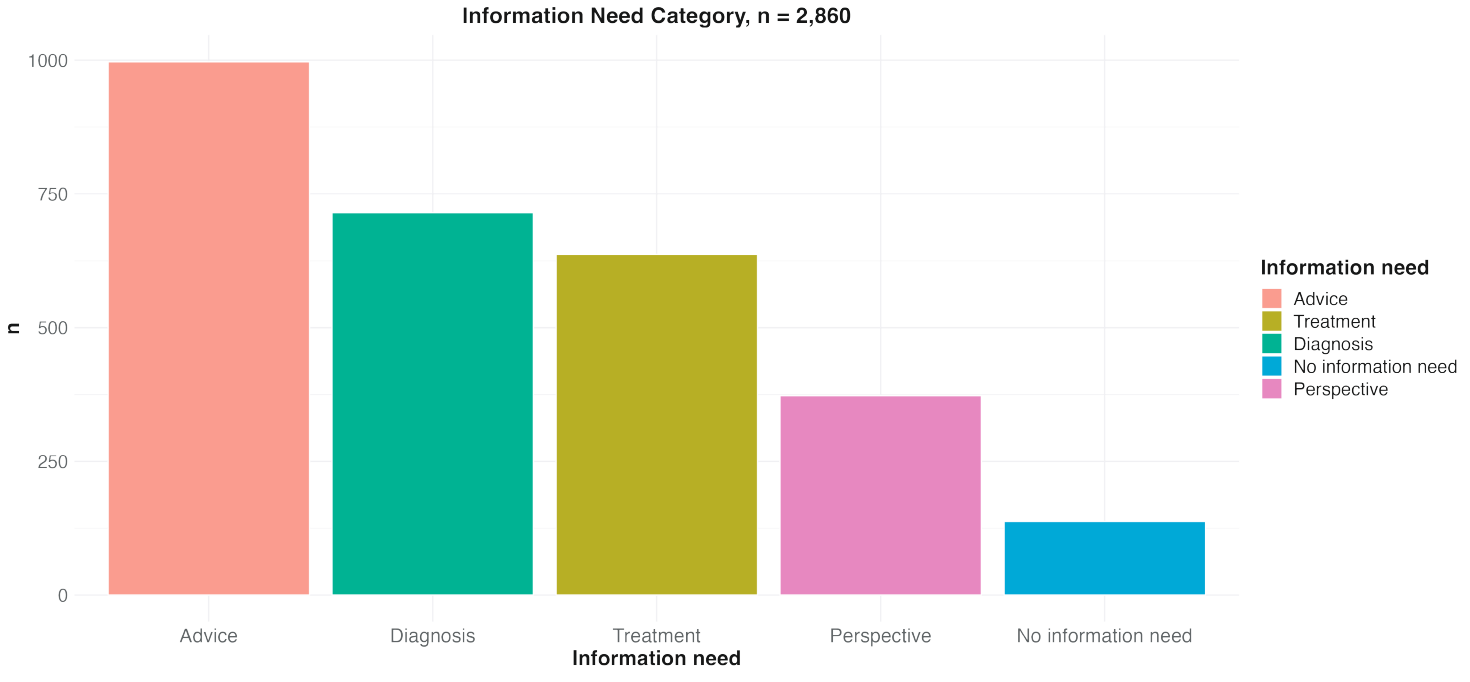


Fig. 3: Distribution of the information need in the final version of the Patient-Specific Needs Evaluation (PSN). The patient chooses one of the following options: I do not need information; Diagnosis (I have questions about the diagnosis); Advice (I want to know what is the best thing to do in my situation); Treatment (I have questions about the treatment); Perspective (I want to know what to expect in the future).

Fig. 4: Goal domains chosen as most important in the final version of the Patient-Specific Needs Evaluation (PSN).

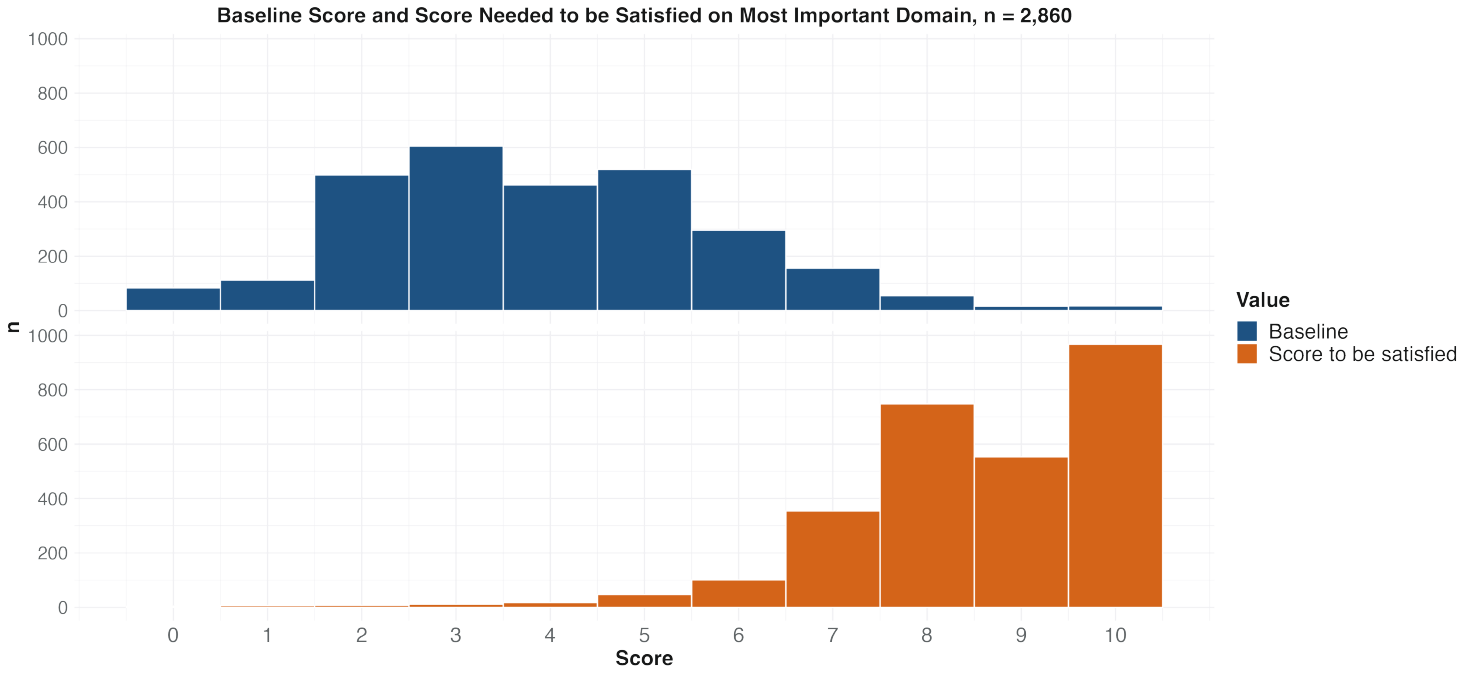


Fig. 5: Distribution of the score on the most important goal domain at baseline and the score patient reported that they needed to achieve to be satisfied with the treatment results in the final version of the Patient-Specific Needs Evaluation (PSN). The figures indicate that not all patients want to obtain the maximum score on their most important outcome domain to become satisfied with the treatment results. The median score needed to be satisfied with the treatment result was 9 in the final version. For ease of interpretation, we converted each domain score to the same scale (i.e., reversing the scores on the domains pain, numbness, and tingling).

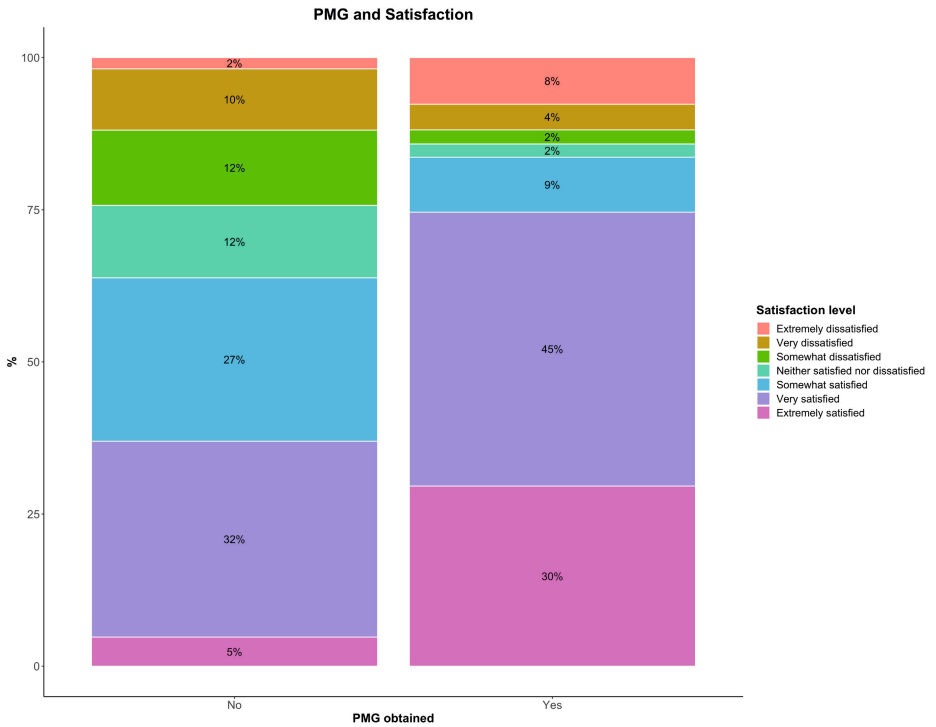


Fig. 6: Discriminative validity of the Personal Meaningful Gain (PMG) in n = 1985 patients, demonstrating that patients that obtained their PMG were much more often satisfied with their treatment results compared to patients who did not obtain their PMG (with a medium to large effect size (Cramer's V: 0.48, p<0.001).

Table 3. Test-retest reliability of the Patient-Specific Needs Questionnaire (PSN).

Test-retest variable	Absolute agreement	Cohen's Kappa (95% CI)	ICC (95% CI)	Conclusion
Most important goal domain	58%	0.46 (0.34 to 0.58)	0.53 (0.38 to 0.66)	Moderate agreement and reliability
Most important goal domain chosen as most important goal domain or as secondary goal domain at retest	75%	0.68 (0.58 to 0.79)	0.73 (0.62 to 0.81)	Substantial agreement and high reliability
Baseline score on most important goal domain	-	-	0.57 (0.42 to 0.69)	Moderate reliability
Score needed to be satisfied at most important goal domain	-	-	0.64 (0.51 to 0.74)	Moderate reliability
Personal Meaningful Gain (PMG) at most important goal domain	-	-	0.65 (0.53 to 0.75)	Moderate reliability

Discussion

The Patient-Specific Needs Evaluation (PSN) focuses on patient-specific information needs and treatment goals and supports patient-centered care. Although developed in hand and wrist patients, the PSN can be modified easily to unlock its potential for generalization by altering answer options. As part of the PSN, we introduce the Personal Meaningful Gain (PMG) as a valid parameter of the improvement an individual wants to obtain in a domain relevant to that individual, given the pre-treatment score.

How to use the PSN

The PSN can be used as a conversation starter, decision-support tool, and expectation management tool during the first consultation. The information needs part facilitates clinicians to effectively provide information and tailor information provision to the individual patient, e.g., knowing a patient's tendency towards surgery may guide how a clinician proposes non-invasive treatment when more appropriate. The treatment goal aids realistic goal setting, e.g., if a patient with Dupuytren's disease wants to improve the hand appearance, but it is unlikely that this will be achieved with treatment. The PMG helps to identify and discuss expectations, e.g., if one wants to improve from 2 to 10 to be satisfied, while this may be unrealistic due to comorbidity or symptom duration. The PSN also evaluates treatment success at a personal level.

There was moderate agreement and reliability for the most important goal domain. However, these improved to a substantial agreement and high reliability when also considering agreement if the most important goal domain was a secondary goal domain in the retest. This indicates that the PSN's reliability is good enough to identify all patient-relevant goals. Thus, patients find it hard to distinguish between their most important and secondary goals, which may overlap. Our finding that most patients who obtained their PMG were satisfied with their treatment results suggests that their satisfaction was independent of whether their PMG was on their factual primary goal, confirming the PSN's useability. Clinicians should always consider all goals, and not only the most important goal domain.

Key considerations

User participation during the development, the iterative approach, pilot testing, and mixed-methods resulted in a content-valid, discriminative, and reliable patient-centered tool. The PSN was easily implemented, and patients deemed it relevant, complete, and understandable. The PSN helps patients prepare for their first consultation, enhances awareness, empowers them to take control of their treatment, and aids shared decision-making. The clinicians indicated that the PSN helps them to identify patients with high or low expectations and respond accordingly. These aspects may improve patients' experience, expectation management, satisfaction with treatment results, and clinical outcomes[51].

Compared to existing tools[33-36], the PSN adds value. For example, the COPM, GAS, and PSGM are completed together with a healthcare provider. Therefore, they are relatively time-consuming in clinical practice, and there is a risk of "therapist bias" as a practitioner may influence these goals. Other tools do not assess patient-specific improvement goals and their relation with satisfaction with treatment results, while the PSN does (i.e., the PMG). Furthermore, in contrast with current tools such as the PSFS, COPM, and PSGM, the PSN allows distinct ICF domains and not only focuses on the activities and participation levels. None of the aforementioned tools assesses information needs, while the PSN does measure these. Altogether, the PSN is a unique tool with added value in daily clinic and research.

The distribution of the information need category and the goal domain indicates that patients have different needs and goals. This highlights that a personalized treatment strategy is essential, which can be informed by the PSN. Further, although most people wanted to reach a 9 to be satisfied, many patients consider lower scores satisfactory, i.e., not all patients aim for the maximum score. The wide distribution indicates that this is indeed a personalized score, which further adds to the value of the PSN.

The PMG distinguished satisfied patients from dissatisfied patients very well, indicating that it can be used to evaluate the clinical relevance of treatment effects. The PMG is especially beneficial as it is determined before clinician consultation, providing a proxy for satisfaction with treatment results at a very early stage, presuming what patients think they want is what they will be satisfied with. Future research may investigate whether the PMG has a greater discriminative capacity for satisfaction than traditional values such as the Minimal Important Change or Patient Acceptable Symptom State.

At our sites, a clinician dashboard is used, which displays, e.g., patient characteristics, PROMs, clinician-reported outcomes (e.g., goniometry), and prediction models. With the PSN added, healthcare can be further personalized and data-driven. Nevertheless, the PSN is also valuable as a stand-alone tool.

We distribute the PSN before surgeon consultation. If treatment is scheduled (e.g., surgery or therapy), we allow patients to change previous answers. For example, the patient's goal may have changed following expectation management during consultation. This strategy is, of course, optional.

Limitations

Respondents indicate their most important needs and goals without knowing their diagnosis. It may also be difficult for individuals to accurately predict how a future score would feel, such as a 9 or 10, since this is an abstract idea that may not match their actual experience when they reach that level. However, focusing on the patient's most important needs and goals at this early stage benefits clinicians, as they may use these in decision-

making and expectation management. Although some items may be moving targets (i.e., a response shift: goals may change over time), the PSN discriminated effectively between satisfied and dissatisfied patients. Future research could investigate how needs and goals change over time. Also, the PSN does not replace traditional outcome measures, and additional time investment should be considered when using it.

Another limitation is the test-retest non-response. The small differences between responders and nonresponders seem clinically irrelevant, as age and type of work are unlikely to influence test-retest reliability. Still, although inevitable in test-retest studies, this may have influenced our findings.

We addressed most issues mentioned by respondents but kept the maximum number of information need categories respondents could choose. Obviously, patients have more questions, and clinicians should try to answer them all. However, we considered it essential that, at least, the most important question is identified and answered as there is a maximum information load persons absorb. Therefore, it is essential to see the PSN as a conversation starter. Also, patients might be better prepared by knowing their most important question[51].

Another limitation is that we excluded patients with inconsistent answers on the survey. This may have influenced our findings on the understandability of the PSN. However, if we had included these patients, our findings would also have been biased; thus, we believe that our decision was the best solution to minimize bias. Also, although the participants had different educational levels (including lower levels), it remains challenging to reach lower literacy patients. Future research may specifically target these.

Although we performed a cross-cultural translation to English, we only tested the Dutch version. Future studies may investigate the PSN in different languages and cultural settings.

Conclusion

The PSN is a novel, brief patient-reported tool identifying individual patient needs and goals. By identifying these, clinicians are better equipped to tailor information provision and treatment to the individual patient, enhancing the quality of care. The PSN can help patients to take control of their treatment. It is valid, reliable, and easy to use, especially but not only in digital form. The PSN is implementation-ready for hand and wrist care and can easily be generalized to other fields. The PSN is provided with open access and is free to use.

References

1. Porter ME. What is value in health care? N Engl J Med. 2010;363(26):2477-81.
2. Barry MJ, Edgman-Levitan S. Shared Decision Making — The Pinnacle of Patient-Centered Care. New Engl J Med. 2012;366(9):780-1.
3. Basch E. Patient-Reported Outcomes — Harnessing Patients' Voices to Improve Clinical Care. New Engl J Med. 2017;376(2):105-8.
4. Bernstein DN, Calfee RP, Hammert WC, Rozental TD, Witkowski ML, Porter ME. Value-Based Health Care in Hand Surgery: Where Are We & Where Do We Go From Here? The Journal of Hand Surgery. 2022.
5. Institute of Medicine Committee on Quality of Health Care in A. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press (US) Copyright 2001 by the National Academy of Sciences. All rights reserved.; 2001.
6. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. N Engl J Med. 2012;366(9):780-1.
7. Basch E. Patient-Reported Outcomes - Harnessing Patients' Voices to Improve Clinical Care. N Engl J Med. 2017;376(2):105-8.
8. Bleich SN, Ozaltin E, Murray CK. How does satisfaction with the health-care system relate to patient experience? Bull World Health Organ. 2009;87(4):271-8.
9. Graham B. Defining and Measuring Patient Satisfaction. J Hand Surg (USA). 2016;41(9):929-31.
10. Marks M, Herren DB, Vliet Vlieland TPM, Simmen BR, Angst F, Goldhahn J. Determinants of patient satisfaction after orthopedic interventions to the hand: A review of the literature. J Hand Ther. 2011;24(4):303-12.
11. De Ridder WA, van Kooij YE, Vermeulen GM, Slijper HP, Selles RW, Wouters RM. Test-retest Reliability and Construct Validity of the Satisfaction with Treatment Result Questionnaire in Patients with Hand and Wrist Conditions: A Prospective Study. Clin Orthop Relat Res. 2021;479(9):2022-32.
12. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, Hovius SER. A Standard Set for Outcome Measurement in Patients With Hand and Wrist Conditions: Consensus by the International Consortium for Health Outcomes Measurement Hand and Wrist Working Group. J Hand Surg Am. 2021;46(10):841-55.e7.
13. De Ridder WA, Wouters RM, Hoogendam L, Vermeulen GM, Slijper HP, Selles RW. Which factors are associated with satisfaction with treatment results in patients with hand and wrist conditions? A large cohort analysis. Clin Orthop Relat Res. 2022.
14. Johnson JE, Nail LM, Lauver D, King KB, Keys H. Reducing the negative impact of radiation therapy on functional status. Cancer. 1988;61(1):46-51.
15. Geissler J, Sharf G, Bombaci F, Daban M, De Jong J, Gavin T, Pelouchova J, Dziwinski E, Hasford J, Hoffmann VS. Factors influencing adherence in CML and ways to improvement: Results of a patient-driven survey of 2546 patients in 63 countries. J Cancer Res Clin Oncol. 2017;143(7):1167-76.

16. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. BMJ Qual Saf. 2014;23(7):534-42.

17. Mandl LA, Burke FD, Shaw Wilgis EF, Lyman S, Katz JN, Chung KC. Could preoperative preferences and expectations influence surgical decision making? Rheumatoid arthritis patients contemplating metacarpophalangeal joint arthroplasty. Plast Reconstr Surg. 2008;121(1):175-80.

18. Poelstra R, Selles RW, Slijper HP, van der Oest MJW, Feitz R, Hovius SER, Porsius JT. Better patients' treatment experiences are associated with better postoperative results in Dupuytren's disease. J Hand Surg Eur Vol. 2018;43(8):848-54.

19. Schrier V, Poelstra R, Selles RW, Slijper HP, Amadio PC, Hovius SER, Porsius JT. Better Patient-Reported Experiences with Health Care Are Associated with Improved Clinical Outcome after Carpal Tunnel Release Surgery. Plast Reconstr Surg. 2019;143(6):1677-84.

20. Tsehaie J, van der Oest MJW, Poelstra R, Selles RW, Feitz R, Slijper HP, Hovius SER, Porsius JT. Positive experience with treatment is associated with better surgical outcome in trapeziometacarpal osteoarthritis. J Hand Surg Eur Vol. 2019;44(7):714-21.

21. Selles RW, Wouters RM, Poelstra R, van der Oest MJW, Porsius JT, Hovius SER, Moojen TM, van Kooij Y, Pennehouat PY, van Huis R, Vermeulen GM, Feitz R, Slijper HP. Routine health outcome measurement: development, design, and implementation of the Hand and Wrist Cohort. Plast Reconstr Surg. 2020;146(2):343-54.

22. Leventhal HLBI, Brissette, I., & Leventhal, E. A. The common-sense model of self-regulation of health and illness. The self-regulation of health and illness behaviour. 12003. p. 42-65.

23. Husson O, Mols F, van de Poll-Franse LV. The relation between information provision and health-related quality of life, anxiety and depression among cancer survivors: a systematic review. Ann Oncol. 2011;22(4):761-72.

24. Ormandy P. Defining information need in health - assimilating complex theories derived from information science. Health Expect. 2011;14(1):92-104.

25. Larson CO, Nelson EC, Gustafson D, Batalden PB. The relationship between meeting patients' information needs and their satisfaction with hospital care and general health status outcomes. Int J Qual Health Care. 1996;8(5):447-56.

26. Husson O, Thong MS, Mols F, Oerlemans S, Kaptein AA, van de Poll-Franse LV. Illness perceptions in cancer survivors: what is the role of information provision? Psychooncology. 2013;22(3):490-8.

27. Iskandarsyah A, de Klerk C, Suardi DR, Soemitro MP, Sadarjoen SS, Passchier J. Satisfaction with information and its association with illness perception and quality of life in Indonesian breast cancer patients. Support Care Cancer. 2013;21(11):2999-3007.

28. Wahl AK, Robinson HS, Langeland E, Larsen MH, Krogstad AL, Moum T. Clinical characteristics associated with illness perception in psoriasis. Acta Derm Venereol. 2014;94(3):271-5.

29. Hagger MS, Koch S, Chatzisarantis NLD, Orbell S. The common sense model of self-regulation: Meta-analysis and test of a process model. Psychol Bull. 2017;143(11):1117-54.

30. Hazard RG, Spratt KF, McDonough CM, Carayannopoulos AG, Olson CM, Reeves V, Sperry ML, Ossen ES. The impact of personal functional goal achievement on patient satisfaction with progress one year following completion of a functional restoration program for chronic disabling spinal disorders. Spine (Phila Pa 1976). 2009;34(25):2797-802.

31. Levack WM, Taylor K, Siegert RJ, Dean SG, McPherson KM, Weatherall M. Is goal planning in rehabilitation effective? A systematic review. Clin Rehabil. 2006;20(9):739-55.

32. Stevens A, Beurskens A, Köke A, van der Weijden T. The use of patient-specific measurement instruments in the process of goal-setting: a systematic review of available instruments and their feasibility. Clin Rehabil. 2013;27(11):1005-19.

33. Law M, Baptiste S, McColl M, Opzoomer A, Polatajko H, Pollock N. The Canadian occupational performance measure: an outcome measure for occupational therapy. Canadian Journal of Occupational Therapy. 1990;57(2):82-7.

34. Kiresuk TJ, Sherman RE. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Ment Health J. 1968;4(6):443-53.

35. Stevens A, Köke A, van der Weijden T, Beurskens A. The development of a patient-specific method for physiotherapy goal setting: a user-centered design. Disability and Rehabilitation. 2018;40(17):2048-55.

36. Stratford PW GC, Westaway M, Binkley JM. Assessing disability and change on individual patients: a report of a patient-specific measure. Physiotherapv Canada. 1995(47):258-63.

37. Organization WH. International Classification of Functioning, Disability and Health (ICF) Geneve2001 [Available from: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health.

38. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res. 2021;30(8):2197-218.

39. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. J Orthop Trauma. 1998;12(8):577-86.

40. MacDermid JC, Grewal R. Development and validation of the patient-rated ulnar nerve evaluation. BMC Musculoskelet Disord. 2013;14:146.

41. de Kraker M, Selles RW, Schreuders TA, Hovius SE, Stam HJ. The Pollexograph: a new device for palmar abduction measurements of the thumb. J Hand Ther. 2009;22(3):271-6; quiz 7.

42. Schreuders TA, Selles RW, Roebroeck ME, Stam HJ. Strength measurements of the intrinsic hand muscles: a review of the development and evaluation of the Rotterdam intrinsic hand myometer. J Hand Ther. 2006;19(4):393-401; quiz 2.

43. Feitz R, van Kooij YE, Ter Stege MHP, van der Oest MJW, Souer JS, Wouters RM, Slijper HP, Selles RW, Hovius SER. Closing the loop: a 10-year experience with routine outcome measurements to improve treatment in hand surgery. EFORT Open Rev. 2021;6(6):439-50.

44. de Vet HCW, Terwee, Caroline B., Mokkink, Lidwine B., Knol DL. Measurement in Medicine. Cambridge: Cambridge University Press; 2011.

45. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.

46. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968;70(4):213-20.

47. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.

48. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420-8.

49. Mitani AA, Freer PE, Nelson KP. Summary measures of agreement and association between many raters' ordinal classifications. Ann Epidemiol. 2017;27(10):677-85.e4.

50. Fleiss JL. Reliability of Measurement. The Design and Analysis of Clinical Experiments1999. p. 1-32.

51. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res. 2009;18(1):115-23.

Supplemental content 1

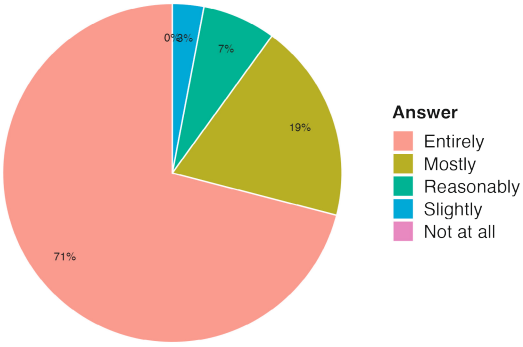
Main categories	Findings	Quotes
1: Relevance	1.1: Most participants found it relevant that their answers could direct the conversation	1.1.1 "I think it's important for a doctor to know that in advance. Know what the situation is" (P1) 1.1.2 "easy to give direction to the conversation" (P1) 1.1.3 "Time efficient" (P2)
	1.2: Most participants found it relevant that the clinician could respond to expectations	1.2.1 "If you absolutely want to reach a 10, that says something about you" (P1) 1.2.2 "it is easier for the doctor to estimate the patient's expectations" (P3) 1.2.3 "I can imagine that it is nice for the practitioner to already know what someone would most like to improve" (P4)
	1.3: Some participants felt well prepared and empowered by filling in the questionnaire	1.3.1 "Yes, also nice for myself to already think about what I would like to know during such a first intake, for example" (P5) 1.3.2 "You feel that you are taken seriously" (P6)
2: Completeness: information need	2.1 Some participants seem worried about whether the clinician would answer all their questions	2.1.1 FIELD NOTES
	2.2 Many participants deemed it impossible to choose just one answer option	2.2.1 "To me, everything was important" (P2) 2.2.2 "asking multiple questions to arrive at an opinion" (P6)
3: Completeness: treatment goal and improvement goal	3.1 All participants thought the questions and answer options were complete	3.1.1 "I believe everyone can find his or her complaint in here" (P7) 3.1.2 "Seems all right to me, yes, I can't see one missing" (P1)
4: Understandability: introductory texts	4.1 Many participants thought the welcome text was too elaborate	4.1.1 "It could also be said in two sentences" (P7) 4.1.2 "Yes, just a good clear explanation" (P6)
	4.2 Some participants missed a sentence on how many options they were allowed to pick at 'treatment goal'	4.2.1 "It literally says pick one" (P3) 4.2.2 "for some people I think it is annoying to have to focus on one thing again" (P3)
5: Understandability: information need	5.1 Most participants found the question and answer options clear	5.1.1 "Give me a second... No, it's clear like this" (P8)
	5.2 Some participants thought they couldn't answer the question without more information beforehand	5.2.1 "You can't really ask about treatment if a diagnosis has not yet been determined" (P3)
6: Understandability: treatment goal and improvement goal	6.1 Many participants thought the answer scale for pain was too complicated	6.1.1 "You need to switch the scale" (P3) 6.1.2 "With smiley faces for people who think it is complicated" (P6) 6.1.3 "I would say 'unbearable pain' and 'no pain' instead of 'very bad' and 'excellent'" (P3)
	6.2 Some participants found specific words in the domain section too complicated	6.2.1 "Well, I didn't know what pliability was" (P2)

Supplemental content 2

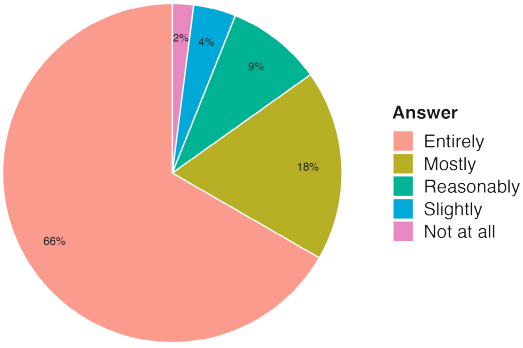
Main categories	Findings	Quotes
1: Relevance	1.1 Many clinicians thought the open question for aid and the conversation itself provide enough information to understand what the patients needs are	1.1.1 "you can just have this conversation and then you will get this information too" (C2) 1.1.2 "I'm quite satisfied with the information that we already have" (C4)
	1.2 Some clinicians thought the questionnaire might help them to formulate a relevant treatment goal	1.2.1 "I would definitely look it over and see what the patients goal was prior to their visit to the doctor. And of course I would discuss that with them, before coming here, your goal was this, is that still your goal or do you now want something else?" (C1)
	1.3 Many clinicians thought it useful to know what the patients goal is to become satisfied with the treatment results	1.3.1 "If the expectations aren't realistic then I would use it. (...) If it is just regular, then I won't do much with it" (C4) 1.3.2 "you can filter out those extremes nicely"(C2)
	1.4 Some clinicians thought it useful that the patient fills in the questionnaire before the first consultation	1.4.1 "I actually think this is a good one, because the patient can tell his own story. So one is a little less likely to be overwhelmed by the opinion of a clinician" (C3)
	1.5 Some clinicians thought it useful to have an image of the patient before the first consultation	1.5.1 "It's interesting, by certain answers, you also get to see a kind of personality" (C5)
2: Completeness: information need	2.1 All clinicians think the part on information need is complete	2.1.1 "It's definitely complete, especially the first question" (C1)
3: Completeness: treatment goal and improvement goal	3.1 Most clinicians think the part on the treatment goal is complete	3.1.1 "Yes, yes, I think it is fairly complete in terms of complaints" (C3)
4: Usability for the clinician	4.1 Barrier: Most clinicians are afraid the questionnaire will cost them more time	4.1.1 "because of course you don't have forever to prepare so I'm not sure I would look at this" (C2) 4.1.2 "Example given, 'oh yes, I saw that you are a bricklayer or something' and then you immediately have a conversation and someone also has the feeling that his information is used" (C3)
	4.2 Risk: Some clinicians fear being biased	4.2.1 "Well, I think I have to be very careful not to start with prejudices. Someone has discussed his profession and his complaints, so I'm already starting with a tunnel vision" (C3) 4.2.2 "And especially a conversation is dynamic. You can't put at person in a box" (C4)
	4.3 Risk: some clinicians fear that clinicians will only answer the most important answer	4.3.1 "If I look at it quickly, I could just be put on the wrong track when I see that they can only indicate one" 4.3.2 "You have to let therapists know that patients are only asked to only choose one answer" (C1)
	4.4 Barrier: Some clinicians thought it hard to use the patients answer to the information need in their consult	4.4.1 "And vice versa, you choose diagnosis and treatment, or advice and future, but that is usually also a multi-question" (C4) 4.4.2 "Yeah, I honestly don't know if I'll be using this when they can only choose one option, because then I know okay, they will ask more questions anyway. Do you understand what I mean?" (C2)
5: Usability/ understandability for patients	5.1 All clinicians thought the answer scale to pain was too complicated	5.1.1 "If you just use the same scale for everything and not a grade, because now I will be thinking, 'hey, a grade, okay, but I just entered something different; help, did I do that right?'" (C2)
	5.2 Some clinicians think several words and questions are too hard for patients to understand	5.2.1 FIELD NOTE: suggestion to put answers in sentences (narrative mode) (C4, C5, C6) 5.2.2 "We have a certain level of intelligence, it's not that I feel elevated, but a majority of patients do not even understand some of the words" (C3)
	5.3 Some clinicians wonder whether the patient answers honestly	5.3.1 "All patients want their doctor to put maximum effort in it" (C5) 5.3.2 "But is that realistic?" (C1)
	5.4 Risk: many clinicians are afraid that patients have to answer too many overlapping questions	5.4.1 "Yes, you know, whatever, 'I already filled this in' and then you fill in the question differently than the other one. So then you give a score, you don't look at it carefully, while you might have done that other list very carefully" (C3)

Supplemental content 3

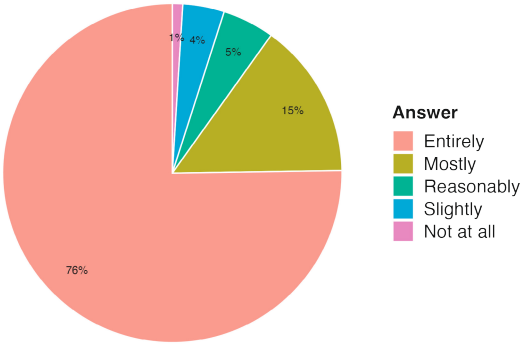
A **Did you find the question about your information need understandable?**



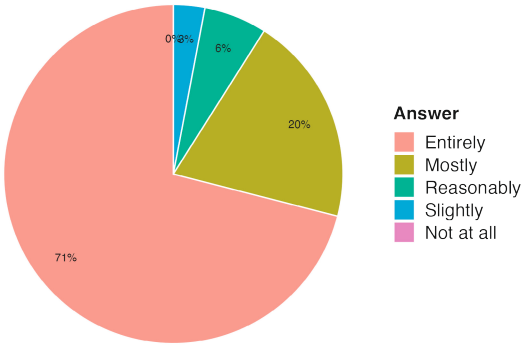
C **Are the response options on your information need complete?**



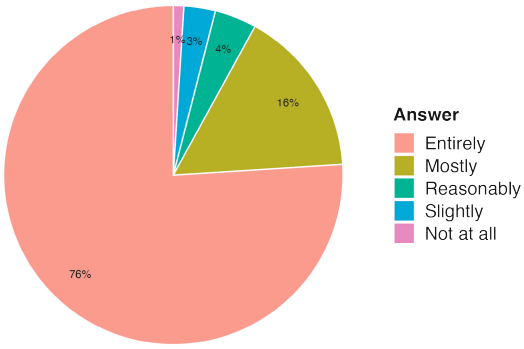
E **Did you find the response options on your treatment goals understandable?**



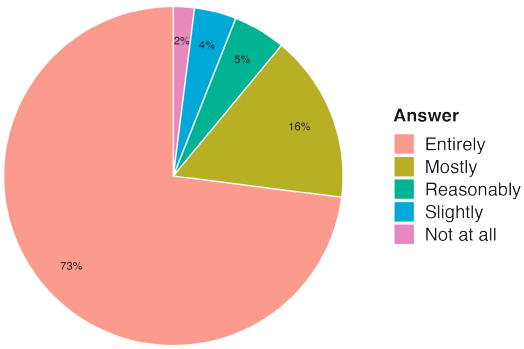
B **Did you find the response options on your information need understandable?**



D **Were the questions about your treatment goals understandable?**



F **Are the response options on your treatment goals complete?**



Supplemental Digital Content 4. Non-responder analysis for the test-retest study.

Variable	Non-Responders (n = 37)	Responders (n = 102)	p-value*	SMD
Age, median [IQR]	53.00 [36.00, 63.00]	64.00 [51.25, 73.75]	<0.001	0.691
Sex = male, n (%)	13 (35.1)	46 (45.1)	0.392	0.204
Duration of symptoms in months, median [IQR]	12.00 [6.00, 28.00]	11.00 [5.00, 18.75]	0.171	0.305
Type of work, n (%)			0.009	0.745
Unemployed due retirement	6 (16.2)	41 (40.2)		
Unemployed due other reason	4 (10.8)	6 (5.9)		
Light physical labor (e.g., office work)	10 (27.0)	22 (21.6)		
Moderate physical labor (e.g., working in a store)	14 (37.8)	16 (15.7)		
Heavy physical labor (e.g., working in construction	3 (8.1)	17 (16.7)		
Level of education (%)			0.950	0.328
None	1 (2.7)	1 (1.0)		
Primary education (primary school, special primary education)	0 (0.0)	1 (1.0)		
Primary or pre-vocational education (such as (in Dutch) LTS, LEAO, LHNO, Huishoudschool, VMBO)	4 (10.8)	12 (11.8)		
Secondary general secondary education (such as (in Dutch) MAVO, (M)ULO, MBO-short, VMBO-t)	6 (16.2)	24 (23.5)		
Secondary vocational education and vocational training (such as (in Dutch) MKBO-long, MTS, MEAO, BOL, BBL, INAS)	8 (21.6)	20 (19.6)		
Higher general and pre-university education (such as (in Dutch) HAVO, VWO, Atheneum, Gymnasium, HBS, MMS)	2 (5.4)	9 (8.8)		
Higher vocational education (such as (in Dutch) HBO, HTS, HEAO, HBO-V, university graduates	9 (24.3)	21 (20.6)		

Supplemental Digital Content 4. Non-responder analysis for the test-retest study. *(continued)*

Variable	Non-Responders (n = 37)	Responders (n = 102)	p-value*	SMD
Scientific education (e.g., MSc.)	4 (10.8)	8 (7.8)		
Prefer not to say	3 (8.1)	6 (5.9)		
Body Mass Index, median [IQR]	26.00 [23.00, 28.00]	26.00 [23.00, 29.00]	0.617	0.042
Smoking status, n (%)			0.366	0.356
Yes, daily smoker	7 (18.9)	10 (9.8)		
Yes, passive smoker	0 (0.0)	2 (2.0)		
Yes, sometimes	1 (2.7)	6 (5.9)		
No	29 (78.4)	84 (82.4)		
Affected side, n (%)			0.953	0.059
Left	13 (35.1)	33 (32.4)		
Right	14 (37.8)	40 (39.2)		
Both	10 (27.0)	29 (28.4)		
Dominance, n (%)			0.560	0.223
Left	3 (8.1)	11 (10.8)		
Right	33 (89.2)	84 (82.4)		
Both	1 (2.7)	7 (6.9)		
Second opinion = no, n (%)	34 (91.9)	87 (85.3)	0.460	0.209
Personal injury lawsuit = no, n (%)	36 (97.3)	100 (98.0)	1.000	0.049

*Continuous variables were compared using the Kruskal-Wallis Rank Sum Test and dichotomous or categorical variables using a Chi-Square test.

Supplemental Digital Content 5. Cross table demonstrating how often the most important goal domain was chosen at the primary test as well as at the retest. The values correspond to the number of patients and the percentage of the row total, except for the "Row total" column, where the percentages correspond to the percentage of the column total.

	Most important goal domain at retest								
Most important goal domain at primary test	*Pain*	*Activities*	*Flexibility/ Mobility*	*Work*	*Tingling*	*Strength*	*Appearance*	*Numbness/ Sensation*	*Row total*
Pain	24 (69%)	5 (0,14%)	3 (9%)	2 (6%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	35 (34%)
Activities	7 (29%)	15 (63%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	24 (24%)
Flexibility/Mobility	2 (13%)	3 (20%)	8 (53%)	0 (0%)	0 (0%)	0 (0%)	2 (13%)	0 (0%)	15 (15%)
Work	1 (13%)	2 (25%)	1 (13%)	3 (38%)	0 (0%)	0 (0%)	0 (0%)	1 (13%)	8 (8%)
Tingling	0 (0%)	1 (17%)	0 (0%)	0 (0%)	5 (8%)	0 (0%)	0 (0%)	0 (0%)	6 (6%)
Strength	2 (33%)	2 (33%)	0 (0%)	0 (0%)	0 (0%)	1 (0,167)	1 (17%)	0 (0%)	6 (6%)
Appearance	0 (0%)	0 (0%)	1 (50%)	0 (0%)	0 (0%)	0 (0%)	1 (50%)	0 (0%)	2 (2%)
Numbness/Sensation	0 (0%)	1 (20%)	1 (20%)	0 (0%)	1 (20%)	0 (0%)	0 (0%)	2 (40%)	5 (5%)
No treatment goal	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)
Column total	36 (35%)	29 (28%)	14 (14%)	6 (6%)	7 (7%)	1 (1%)	5 (5%)	4 (4%)	102 (100%)

7

11

GENERAL DISCUSSION

General discussion

The goal of this thesis was to enhance patient-centered and value-based care by improving satisfaction with treatment results (STR) in patients with hand or wrist disorders. To achieve this, the thesis aimed to:

1. Develop a more comprehensive understanding of satisfaction with treatment results and its related factors in patients with hand or wrist disorders;
2. Explore the connection with the patients' mindset;
3. Improve satisfaction with treatment results using data-driven tools.

This general discussion follows the structure of the thesis: 1. Measure and understand STR; 2. Explore the connection of STR with mental health and outcome expectations; and 3. Improve STR using data-driven tools. First, an exploration of the main findings, implications, and future perspectives is undertaken for each part individually. Following this, I consider the most significant limitations of this thesis. Finally, the discussion concludes with recommendations tailored for clinicians, researchers, and policymakers.

Part 1. Measure and understand satisfaction with treatment results

The aims of Part 1 were:

- To investigate the psychometric properties of measures for evaluating satisfaction with treatment results
- To identify factors associated with satisfaction with treatment results

In Part 1, we investigated the test-retest reliability and construct validity of the Satisfaction with Treatment Result Questionnaire (STRQ). The questionnaire was filled in twice three months after treatment initiation by 174 patients, which we used for calculating the test-retest reliability. For the construct validity, we did hypothesis testing using 3742 patients who completed the STRQ, VAS pain and hand function, and the Net Promotor Score (NPS) at 3 months. We concluded that the STRQ is a reliable and valid tool for evaluating patients' satisfaction with their treatment results after hand or wrist treatments, and can be used in both clinical practice and research.

Knowing that we could safely use the STRQ, we investigated factors explaining STR. The STRQ included two questions: "Are you satisfied with the treatment result so far?" and "Would you be willing to undergo the treatment again under similar circumstances?". Using two logistic hierarchical regression models in 1824 patients with common hand or wrist conditions, we found a very high proportion of the findings explained by the variables in our model, namely 82% of the variation in STR and 81% of the willingness to undergo the treatment again. This indicates an excellent ability to distinguish between satisfied and dissatisfied patients and between patients who are or are not willing to undergo the treatment again. We identified several factors associated with one or both of the STRQ questions, such as, amongst others, a greater decrease in pain following

treatment, the patient's positive experience with the explanation of the pros and cons of the treatment and shared decision-making, higher outcome expectations, and better illness perception.

Implications and future perspectives of Part 1

STR is considered a multi-dimensional and complex construct, leading to previously expressed critiques about its measurement, interpretation, and application[1-4]. Therefore, we started this thesis by investigating the validity of these critiques. One of the most significant expressed doubts was whether STR accurately reflects treatment outcomes or merely measures 'peripheral matters'[1-3], such as the mental health of the patient[5], the experienced empathy of the healthcare provider[6], or the time that the healthcare provider spends with their patient. Based on our findings in Part 1, we conclude that STR can be measured with confidence, we now know what factors influence STR, and what STR reflects, that is, the patient's opinion of the treatment results as a multi-dimensional construct, strongly influenced by the patient's illness perceptions, outcome expectations, and experience with healthcare delivery. Similarly, it's important to recognize that other PROMs, such as assessments of pain, hand function, or overall quality of life, are also multi-dimensional in nature. For instance, when a patient reports their level of pain, it encompasses various aspects like intensity, frequency, and its impact on daily activities. Likewise, evaluations of hand function or quality of life involve a range of factors and experiences. Therefore, the multi-dimensionality of STR is not an isolated challenge; in fact, it's a characteristic shared by many assessments in healthcare. What this emphasizes is the need for a comprehensive and nuanced understanding of these measures. By acknowledging and considering the various dimensions involved, clinicians can more effectively interpret and utilize these assessments in guiding treatment decisions. This broader perspective ensures that the true depth of patient experiences and outcomes is taken into account, leading to more informed and tailored care.

Another point of criticism is that patients may lack the necessary knowledge to evaluate treatment outcomes, e.g., a patient can never know what the outcome would have been if they had received a different treatment or the same treatment carried out by a different clinician. Patients may also have difficulty distinguishing between the effects of the treatment and the natural course of their condition. As a result of both, it has been argued that patients can not know if they should be satisfied. While this indeed may play a role in the judgement of patients, it's worth noting that clinicians face similar challenges. While they possess medical expertise, relying solely on personal experience or gut instinct for decision-making is inherently unreliable[7]. Clinicians, too, encounter difficulty in discerning the precise effects of treatment and separating them from the natural course of a condition. This underscores the importance of leveraging real-world data in the decision-making process for both patients and clinicians. Real-world data, such as used for this thesis, provides a comprehensive view of how treatments perform in everyday clinical practice, capturing the nuances and complexities that may not be readily

apparent in controlled research settings. By utilizing this rich source of information, both patients and clinicians can make better-informed decisions regarding treatment options.

In terms of applicability, it is clear that clinicians should prioritize appropriate and effective care over STR when necessary. There may be situations where a patient's desire for a specific treatment option conflicts with the clinician's professional judgment. For example, a patient may insist on surgery even though a less invasive treatment option would be more appropriate. Additionally, there may be cases where a patient is very satisfied with the results of a placebo or ineffective treatment or where a patient is satisfied with the results while the clinician believes a better outcome could have been possible. In these instances, clinicians must balance the patient's wishes and probable future satisfaction with the need to provide the best possible, evidence-based, and valuable care. Given the high healthcare costs in countries like the Netherlands, it is especially important to investigate and offer less invasive, less impactful, and less expensive care that still produces the best possible outcomes, including high STR[8,9]. Therefore, it is imperative for future research to explore and develop strategies that bolster less invasive treatment options.

Based on the studies in Part 1, we now have more insight into the concept of STR and may have resolved some of the doubts about its usage in healthcare. We know that we can measure STR in a reliable and valid manner, and we recommend using the STRQ or a similar tool. Patients think STR is an essential outcome domain[9], and clinicians can use data on STR as an integral part of the evaluation of their treatment and compare scores with other clinics or other clinicians. Moreover, as Part 1 identified several influenceable factors that improve STR, future studies could investigate interventions that do so. Interventions can range from using mandatory checklists for patient information to influence their pre-treatment mindset, providing personalized information electronically or through a chatbot post-consultation. Additional approaches involve psycho-education for clinicians and patients to enhance expectations or understanding, along with decision-support tools and prediction models. We will discuss this more elaborately in Part 2 and 3.

Part 2. Explore the connection with the patients' mindset

The aims of Part 2 were:

- To identify factors associated with pre-treatment outcome expectations
- To evaluate the change in mental health following the first hand surgeon consultation

To gain more insight into mental health and outcome expectations as important concepts related to satisfaction with treatment results, we aimed to identify factors associated with pre-treatment outcome expectations in patients with hand and wrist conditions. Through a cross-sectional study involving 12,345 patients, using a multi-level hierarchical regression model, we found that outcome expectations were primarily determined by the invasiveness of the treatment and patients' illness perceptions. Patients scheduled for

minor or major surgery had higher outcome expectations compared to those scheduled for less invasive treatment, while patients who expected a longer illness duration and were treated for the same condition before had lower outcome expectations.

To examine the impact of the first consultation with a hand surgeon on mental health, in Chapter 5, we evaluated changes in illness perception, psychological distress, and pain catastrophizing following the first surgeon consultation. Our results showed that the total score and almost all subscales of illness perception, as measured by the Brief Illness Perception Questionnaire (B-IPQ), improved after consultation. Additionally, we found that surgical patients improve more compared to non-surgical patients. Last but not least, patients also had decreased levels of pain catastrophizing.

Implications and future perspectives of Part 2

While there is increasing evidence supporting the relationship between more positive expectations and better outcomes in healthcare, some authors caution that patients may already have excessively high expectations towards medical treatments, and suggest that clinicians should help manage, or temper, these expectations[10-13]. One potential downside of having elevated expectations for treatment is an increased likelihood of post-treatment dissatisfaction with treatment results if these expectations are not met. This may be due to a variety of reasons, such as unrealistic expectations or a negative mindset that makes it difficult to appreciate the benefits of the treatment.

Since higher expectations lead to better outcomes and higher STR, but unrealistic high expectations may lead to the opposite, we propose an individualized approach, where high though realistic expectations are optimum. In such an approach, first of all, information provision on the different treatment options, possible outcomes from each option, and associated limitations will help patients make informed decisions about their treatment and help set realistic expectations for treatment outcomes. Second, to actively manage expectations, the clinician needs to know three elements: whether the patient has a more positive or a more negative mindset, has more positive or negative expectations, and whether these expectations are realistic or not. This results in a preliminary model in which we could classify patients into four categories: patients with 1) a negative mindset and realistic expectations; 2) a negative mindset and unrealistic expectations; 3) a positive mindset and realistic expectations; 4) a positive mindset and unrealistic expectations. For all of these categories, we propose a specific frequently used conversation technique (Figure 1).

Figure 1. Proposed conversation techniques to manage expectations in different types of mindset (negative/positive) in combination with realistic or unrealistic expectations.

1. In case of negative, realistic expectations, clinicians may help the patient to reframe their expectations, by replacing negative words with positive words and by shifting the attention from the negative.
2. In case of negative, unrealistic expectations, clinicians may help the patient to question and change their expectations, e.g., by listing evidence for both the negative expectation and its opposite.
3. In case of positive, realistic expectations, clinicians may enhance the patient's expectations to optimize the self-fulfilling prophecy: expecting good results leads to good results[14].
4. In case of positive, unrealistic expectations, clinicians may stimulate the patient to question and weigh the pros and cons of the treatment.
1. For all these groups, these techniques hypothetically may improve expectations.

Based on these insights, interventions can be developed to help clinicians use information on the patient's mindset in their consultations, e.g., by developing decision-support tools based on patient-reported expectations. It could be useful to implement these skills (more) in their education and guidelines. Also, patients could be empowered by changing their mindset themselves, e.g., by the use of psychoeducation. Future studies may investigate whether this approach is effective.

Furthermore, we found that patients scheduled for surgical treatment have higher expectations if they perceive less personal control over their illness. Hypothetically, these patients may believe that the success of the surgery is largely dependent on the skill of the surgeon and other external factors, and therefore have higher expectations for treatment outcomes if they perceive these external factors to be favorable. Alternatively, patients scheduled for less invasive treatment have higher expectations if they perceive

more personal control over their illness. Similarly, non-surgical patients may believe that the success of the treatment depends more on themselves, and they may be less prone to give the control away to a surgeon. This may be a self-fulfilling prophecy: patients who believe that they control their own recovery may have more treatment adherence and may in turn achieve better results of less invasive treatments. Enhancing personal control, which is indicative of an internal locus of control, is thus beneficial and could potentially enhance treatment adherence, motivation, and self-efficacy, and consequently lead to improved treatment results. In theory, enhancing personal control may also serve as a preventive measure against unnecessary surgical interventions, as patients scheduled for surgery have less personal control than patients scheduled for less invasive treatment[15]. Future studies may focus on interventions to influence personal control, e.g., by focusing on the patient's individual treatment goals, and thereby empowering the patient to reach their goals. Furthermore, researchers could investigate interventions to boost expectations of less invasive treatment.

Lastly, our findings in **Chapter 5** highlight the significance of the first consultation and its possible impact on the patient's mindset. We found a small improvement in the illness perceptions of the patient. Looking at specific parts of the B-IPQ, like Coherence or Concern, I think the improvement could have been much better. For instance, the question under 'Coherence' is about how well the patient understands their illness. By providing information and explaining things during the consultation, the improvement in this area could be more significant. The question under 'Concern' measures how worried the patient is about their illness. Talking about risks, emotions, and concerns during the consultation should ideally help improve this aspect more noticeably. Future research should confirm our findings by doing an experiment, in which the consults are standardized.

Concluding, improving the patient's mindset may lead to better outcomes and better treatment decisions, including choosing more often for less invasive treatment. It would be valuable to develop interventions to increase the effects we found. For example, a guideline or training for clinicians on how to improve outcome expectations or illness perception could be developed and formally tested in an experiment. Also, e-learnings for patients could be used to improve their mindset.

Part 3. Improve satisfaction with treatment results using data-driven tools

The aim of Part 3 was:

- To develop and evaluate tools that help clinicians during daily clinical care to positively respond to each individual patient's mental health, personal information needs, treatment goals, and desired improvements to improve satisfaction with treatment results

In Part 3, we described how we developed and successfully implemented two data-driven tools to improve satisfaction with treatment results. First, we developed the Ultra-Short

Mental Health Screening Tool using 19,156 patients with hand and wrist conditions. The tool consists of 4 items from either the Pain Catastrophizing Scale, the 4-item Patient Health Questionnaire, and the Brief Illness Perception Questionnaire and was found to have high construct and discriminative validity and high test-retest reliability. The median response time for patients to fill in the screener was 43 seconds compared to over 4 minutes for the full questionnaires. While maintaining a high validity, the time gain reduces the burden of the PROMs patients are asked to fill in.

Another tool we developed is the Patient-Specific Needs evaluation (PSN). The PSN evaluates individual information needs, treatment goals, and a novel concept of Personal Meaningful Gain (PMG). We developed the PSN to ensure alignment between patient and clinician needs and goals. Moreover, we deemed it necessary to develop an individually tailored outcome measure, precisely measuring what is important to the individual rather than what is important to the average patient. The PMG measures the minimum improvement relevant to the patient to be satisfied with the treatment result. Our study showed that patients who reached their PMG were more satisfied with the results than patients who didn't reach their PMG, which confirms the good performance of the PMG.

Many additional research questions arise from the developing PSN, some of which we answered in the remaining chapters of this thesis. Using the information need part of the PSN, in Chapter 8 we examined the associations between sociodemographics, mental health and expectations, treatment type, and patient-reported outcomes with the fulfillment of information needs and the experience with information provision. Results showed that 66% of patients rated the fulfillment of their information needs at 8 or higher (range: 0-10, 10 is completely fulfilled) three months after treatment. For both the fulfillment of information needs and the experience with information provision, we found the strongest association with the patient's mindset.

Then, we dived deeper into the properties and potential of the PMG in Chapters 9 and 10. We evaluated the PMG in 5133 patients treated for one of four common hand or wrist conditions, stratified by goal domain, and compared surgical and nonsurgical patients. Surgical patients had higher PMGs (i.e., more ambitious treatment and improvement goals) than nonsurgical patients, irrespective of diagnosis; in other words, surgical patients need a greater improvement to be satisfied with the treatment result than non-surgical patients. The differences we found underline the nuanced relationships between diagnosis, treatment approach, and patient expectations, emphasizing the need for personalized healthcare strategies.

Finally, we compared the PMG with the MIC and PASS of validated PROMs, including the MHQ, the PRWHE, the BCTQ-SSS, and pain during loading, pain in rest, and hand function measured on a Numeric Rating Scale (NRS). The PMG consistently outperformed both PASS and MIC in identifying satisfied patients across analyses. This shows the PMGs

superiority as a clinical outcome threshold for treatment success (provided that treatment success equals a satisfied patient).

Implications and future perspectives of Part 3

In recent years, patient-centered and value-based healthcare models have garnered widespread recognition, putting the patient first and striving for high-quality care at reduced costs. Central to these frameworks is the systematic collection and assessment of patient data, referred to as Patient-Reported Outcome Measures (PROMs). Despite the growing adoption of routine outcome measurement in daily healthcare, the potential to improve the quality of care in individual patients has not yet been fully exploited.

First, PROMs have been carefully developed over the years with thoroughness and precision prioritized over brevity, especially for use in research settings like randomized controlled trials (RCTs). In this thesis, we focus on everyday clinical practice, where the time and effort patients invest in completing assessments become pivotal factors. Therefore, we developed the Ultra-Short Mental Health Screening Tool (**Chapter 6**). With this tool, the patient burden is much less, while it offers the clinician an indication of the patient's mental health. This provides the opportunity to discuss mental health issues and to consider them when making treatment decisions, knowing that better mental health leads to better outcomes. Future research should focus on developing brief versions of PROMs or the development of brief tools that support decision-making in everyday practice.

Second, the assessment of treatment effectiveness has traditionally relied on standardized clinical measures, particularly focused on aspects like pain and function. This limited view of assessing treatment effectiveness does not do justice to the needs of the individual patient. With the development of the Patient-specific Needs evaluation (**Chapter 7**), we redirect our focus toward domains of treatment outcomes that have historically been given little attention in treatment effectiveness evaluation, such as appearance or work ability. In the PSN, the patient's assessment of treatment effectiveness is centered on a domain that is personally relevant and chosen by the patient.

Third, in daily clinical practice, decisions are often based on clinician experience and established treatment guidelines, drawn from aggregated data like outcomes in randomized controlled trials. While this approach is valuable, it may not fully meet individual patient needs. Using a uniform benchmark for all patients, assuming a consistent definition of treatment success or failure, may not adequately consider the unique requirements of individual patients. Current benchmarks typically involve achieving a change in a specific outcome measure greater than the Minimal Clinically Important Difference or reaching the Patient Acceptable Symptom State[16]. While useful for assessing treatment impact on a group level, they may not provide the same level of relevance for individual patients, as the threshold for meaningful improvement depends

on factors like the patient's specific condition, type of treatment, initial score, and other patient-specific considerations. Therefore, we introduced the Personal Meaningful Gain (**Chapter 7**), quantifying the minimum improvement needed for a patient to consider the treatment outcome satisfactory. Initial findings highlight the significant impact of achieving one's PMG on overall satisfaction with treatment results. Since we found that PMG outperforms the PASS and the MIC, actively incorporating the PMG into clinical practice and research can elevate the precision of treatment evaluations. Moreover, the PMG offers an interesting opportunity to be used in real-time prediction models for (cost-)effectiveness at an individual level. This is in line with the value-based healthcare framework and will facilitate shared decision-making and provide greater patient value, yielding higher satisfaction with treatment results and reduced costs.

While these tools represent a promising advancement toward patient-centered care, it is important to acknowledge that implementing such a personalized approach in routine clinical practice presents its own set of challenges. Factors such as integration into existing healthcare systems and workflows, as well as adaptation for different patient populations and clinical settings, and adoption by healthcare professionals, warrant careful consideration and sustained effort.

Limitations of this thesis

Most of the studies in this thesis are based on observational data. Although an observational study design offers the advantage of reflecting real-life clinical practice and large patient samples, a limitation of this approach is non-response. In many chapters, a substantial proportion of patients did not respond, posing a potential limitation to the findings. However, non-responder analyses and additional Little tests strongly suggest that the data were missing at random. These findings provide confidence that non-response did not substantially influence the results. Nevertheless, non-response is a common issue in observational studies, and it is possible that non-response may have impacted the results in unforeseen ways. Efforts to minimize non-response and explore effective strategies to improve response rates should be an integral part of conducting research, especially when working with observational data.

A related limitation concerns our study in **Chapter 5**, where we have found that the patients' mindset improved after the first surgeon consultation. Since we used observational data and we did not interfere with or standardize the consultation, we cannot draw causal conclusions. However, it would have been surprising if the mindset would not have changed. One of the main tasks of the surgeon during this first consultation is to explain, give context, and discuss the patient's concerns, needs, preferences, and values[17]. If anything, we believe that the improvement should have been and could be much larger. Future research could standardize the first consult and test this in an experiment.

Third, in **Chapter 3** we explain satisfaction with treatment results, we included a variety of treatment types, surgical and nonsurgical, which may have led to diluted results due to potential interactions with certain variables. Many studies, also in this thesis, show that patients scheduled for surgery have a different mindset than patients scheduled for less invasive treatment[15,18-20]. Therefore, it would have been interesting to stratify surgical and nonsurgical patients. Nevertheless, by adjusting for treatment type in our analysis and finding a small standardized mean difference between treatment type, our findings may very well generalize to a broader population with hand or wrist conditions, and perhaps even to patients with other musculoskeletal conditions.

Conclusion of this thesis and recommendations

This thesis promotes a more personalized approach to healthcare in different ways. This approach provides useful information and tools to improve:

- treatment outcomes such as STR, e.g., through influencing the individual patient's mindset;
- the patient's experience with care, e.g., with personalized information provision;
- treatment decisions and evaluation, e.g., based on individualized and clinically relevant outcomes.

The tools developed emphasize and prioritize the unique needs and goals of each individual. This not only recognizes the diversity of patient needs but also empowers patients to actively engage in their own care decisions. This thesis reinforces the fundamental principle that the essence of healthcare lies in enriching the lives and overall well-being of patients. It emphasizes the imperative that their individual needs, values, and aspirations should guide every medical decision and intervention.

Recommendations to researchers:

- **Develop and evaluate interventions to obtain realistic and high outcome expectations in patients who do not yet have these. Next, interventions to improve other mindset factors should be developed, such as improving illness perceptions and decreasing pain catastrophizing.** Evaluate if this indeed leads to better outcomes, such as higher STR.
- **Enlarge the effect of the first clinician consultation on the patient's mindset. Formalize the consultation to meet this goal and evaluate it in an experiment.** This could be a relatively easy way to improve the patient's mindset and thereby indirectly improve outcomes.
- **Design rituals and tooling for efficient and personalized information provision.** This may improve the experience and the mindset of the patient, leading to better treatment decisions and outcomes.
- **Develop and evaluate interventions to boost the effect of less invasive treatment.** This may lead to fewer risks and lower burden for the patient compared to surgical treatment and more cost-effective treatments.

- **Investigate more applications of the PMG. E.g., whether we can predict if a person will reach their PMG or not, and if so, whether we can use it to make a decision based on cost-effectiveness.** This may lead to better expectation management, better goal-setting, better decision-making, and more cost-effectiveness.

Recommendations to clinicians:

- **Always discuss the pros and cons of treatment, and advise on how to deal with the complaint at home. Apply shared decision-making based on the individual preferences, needs, and goals of the patient.** Make use of applications if this does not come naturally to you. Just do it. It will improve your treatment decisions, patient's STR, outcomes, and experience.
- **Improve the patient's mindset, especially in patients scheduled for less invasive treatment.** Use individualized conversation techniques to influence negative mindsets, such as reframing. Never temper expectations. Be particularly positive about less invasive treatment, since there is a greater need to do so in many patients. This will improve treatment decisions, outcomes, and STR.
- **Use the PSN evaluation, or a tool alike, to elicit the patient's needs and goals. Use the PMG for goalsetting, expectation management, evaluation of the treatment, and shared decision-making.** This will improve decision-making, treatment outcomes, and treatment effectiveness evaluation.

Recommendations to policymakers:

- **Automate, standardize, and digitize healthcare. Support data-driven decision-support tools and personal clinical important outcome values.** This reduces the administrative burden for clinicians and enhances efficacy, better decision-making, and better treatment effectiveness evaluation.
- **Focus on prevention and create a more positive sentiment around less invasive treatment.** This leads to fewer risks and burdens for the patient and more cost-effective treatments.
- **Let go of the holy RCT as the only evidence for treatment effectiveness.** Observational studies have so many advantages and are a better fit to the world we live in**.**

References

1. Ring D, Leopold SS. Editorial-Measuring Satisfaction: Can It Be Done? *Clin Orthop Relat Res.* 2015;473(10):3071-3073.
2. Graham B. Defining and Measuring Patient Satisfaction. *J Hand Surg (USA).* 2016;41(9):929-931.
3. Graham B, Green A, James M, Katz J, Swiontkowski M. Measuring patient satisfaction in orthopaedic surgery. *J Bone Joint Surg Am.* 2015;97(1):80-84.
4. Wilson CD, Mand D, Ring D, Ramtin S. A Systematic Review of Satisfaction Measures in Hand and Wrist Surgery. *J Hand Surg Am.* 2023;48(1):1-8.
5. Duivenvoorden T, Vissers MM, Verhaar JA, Busschbach JJ, Gosens T, Bloem RM, Bierma-Zeinstra SM, Reijman M. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study. *Osteoarthritis Cartilage.* 2013;21(12):1834-1840.
6. Menendez ME, Chen NC, Mudgal CS, Jupiter JB, Ring D. Physician Empathy as a Driver of Hand Surgery Patient Satisfaction. *J Hand Surg (USA).* 2015;40(9):1860-1865.
7. Kahneman D, Sibony, O., & Sunstein, C. R. *Noise: a flaw in human judgment.* Hachette UK; 2021.
8. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477-2481.
9. Wouters RM, Jobi-Odeneye AO, de la Torre A, Joseph A, Hovius SER. A Standard Set for Outcome Measurement in Patients With Hand and Wrist Conditions: Consensus by the International Consortium for Health Outcomes Measurement Hand and Wrist Working Group. *J Hand Surg Am.* 2021;46(10):841-855.e847.
10. Waljee JF, Chung KC. Commentary regarding "Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis". *J Hand Surg Am.* 2015;40(3):491-492.
11. Marks M, Audigé L, Reissner L, Herren DB, Schindele S, Vliet Vlieland TPM. Determinants of patient satisfaction after surgery or corticosteroid injection for trapeziometacarpal osteoarthritis: Results of a prospective cohort study. *Arch Orthop Trauma Surg.* 2015;135(1):141-147.
12. Frouzakis R, Herren DB, Marks M. Evaluation of expectations and expectation fulfillment in patients treated for trapeziometacarpal osteoarthritis. *J Hand Surg [Am].* 2015;40(3):483-490.
13. Kadzielski J, Malhotra LR, Zurakowski D, Lee SG, Jupiter JB, Ring D. Evaluation of preoperative expectations and patient satisfaction after carpal tunnel release. *J Hand Surg Am.* 2008;33(10):1783-1788.
14. Rosenthal R. The Pygmalion effect and its mediating mechanisms. In: *Improving academic achievement.* Academic Press; 2002:25-36.
15. Wouters RM, Vranceanu AM, Slijper HP, Vermeulen GM, van der Oest MJW, Selles RW, Porsius JT. Patients With Thumb-base Osteoarthritis Scheduled for Surgery Have More Symptoms, Worse Psychological Profile, and Higher Expectations Than Nonsurgical Counterparts: A Large Cohort Analysis. *Clin Orthop Relat Res.* 2019;477(12):2735-2746.
16. Terluin B, Koopman JE, Hoogendam L, Griffiths P, Terwee CB, Bjorner JB. Estimating meaningful thresholds for multi-item questionnaires using item response theory. *Qual Life Res.* 2023;32(6):1819-1830.

17. Institute of Medicine Committee on Quality of Health Care in A. In: *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington (DC): National Academies Press (US) Copyright 2001 by the National Academy of Sciences. All rights reserved.; 2001.

18. de Ridder WA, Hoogendam L, Zeroual F, Slijper HP, Wouters RM, Vermeulen GM, Selles RW, van der Oest MJW. Treatment Invasiveness and Illness Perceptions Are Strongly Associated With Outcome Expectations in Patients Treated for Hand or Wrist Conditions: A Cross-sectional Study. *Clin Orthop Relat Res.* 2023;481(5):994-1005.

19. Hoogendam L, van der Oest MJW, Wouters RM, Andrinopoulou ER, Vermeulen GM, Slijper HP, Porsius JT, Selles RW. Patients with higher treatment outcome expectations are more satisfied with the results of nonoperative treatment for thumb base osteoarthritis: a cohort study. *Arch Phys Med Rehabil.* 2021;102(8):1533-1540.

20. Hoogendam L, Koopman JE, van Kooij YE, Feitz R, Hundepool CA, Zhou C, Slijper HP, Selles RW, Wouters RM. What Are the Minimally Important Changes of Four Commonly Used Patient-reported Outcome Measures for 36 Hand and Wrist Condition-Treatment Combinations? *Clin Orthop Relat Res.* 2022;480(6):1152-1166.


11

12

SUMMARY

Summary

The primary objective of this thesis was to advance patient-centered and value-based healthcare by improving satisfaction with treatment results (STR) for patients with hand or wrist disorders. Aligned with the principles of these frameworks, the overarching aim was to enhance patients' well-being while focusing on the crucial role of individual needs, values, and goals in guiding medical decisions and interventions.

To accomplish this, the following objectives were pursued:
1. Develop a more comprehensive understanding of satisfaction with treatment results and its related factors in patients with hand or wrist disorders;
2. Explore the connection with the patients' mindset;
3. Improve satisfaction with treatment results using data-driven tools.

Part 1. Measure and understand satisfaction with treatment results

Since there was no validated metric for measuring STR in patients with hand or wrist conditions, we investigated the reliability and validity of the Satisfaction with Treatment Results Questionnaire (STRQ) in **Chapter 2**. The STRQ was filled in twice three months after treatment initiation by 174 patients, and 3742 patients completed the STRQ, VAS pain and hand function, and the Net Promotor Score (NPS) at 3 months. We found that the STRQ has good to excellent construct validity and high test-retest reliability and can therefore be used to measure STR.

In **Chapter 3**, we then used the STRQ to identify factors associated with STR three months after treatment initiation. In this prospective cohort study, we included patients who underwent carpal tunnel release, nonsurgical or surgical treatment for thumb-base osteoarthritis, trigger finger release, limited fasciectomy for Dupuytren's contracture, or nonsurgical treatment for midcarpal laxity. We performed a logistic hierarchical regression analysis in a sample of 12,345 patients. Results showed that greater decrease in pain during physical load, a positive experience with the explanation of treatment pros and cons, improvement in hand function, a positive experience with advice for at home, better personal control, positive outcome expectations, longer expected illness duration, fewer perceived symptoms, and less concern about the illness were all associated with more STR. Our results indicate that to enhance patient satisfaction with treatment, healthcare providers could improve the patient experience with the healthcare process, influence their perception of the illness, and increase their confidence in the treatment and outcome expectations of the treatment.

Part 2. Explore the connection with the patient's mindset

In Part 2, we explored the connection of the patient's mindset with STR, specifically the domains of outcome expectations and mental health. More positive outcome expectations

have been linked to better treatment outcomes in multiple studies. However, the factors associated with these expectations in clinical practice are not well understood, including patient characteristics and contextual factors like treatment. Studying these factors could provide valuable information for improving outcome expectations and, potentially, treatment outcomes such as STR. Thus, in **Chapter 4**, we investigated which factors determined pre-treatment outcome expectations in patients with hand or wrist conditions. In this cross-sectional study, performing a multi-level linear hierarchical regression analysis including 12,345 patients, we found that more invasive treatment and better illness perceptions are the main factors associated with higher outcome expectations. The outcomes suggest that expectation management should be tailored to the specific treatment (such as surgical versus non-invasive) and the specific patient, including their perception of their illness. More specifically, it may be beneficial to focus on expectation management strategies for non-invasive treatments, such as hand therapy, since these patients have lower expectations.

The patient's mental health plays a crucial role in determining treatment choices and outcomes, such as STR. Improving mental health before treatment could lead to better treatment decisions and better outcomes. Therefore, in **Chapter 5,** we assessed changes in patients' illness perception, psychological distress, and pain catastrophizing following a hand surgeon consultation. In this prospective cohort study, 276 patients with various hand and wrist conditions completed questionnaires before and after the first consultation. The results showed that the overall illness perception, which refers to the thoughts, beliefs, and attitudes that a person holds about their health condition, improved following consultation. Additionally, almost all separate aspects of illness perception improved, and this improvement was greater in patients scheduled for surgical treatment compared to those scheduled for nonsurgical treatment. Furthermore, we found a decrease in pain catastrophizing following the consultation. The improvement of illness perception and pain catastrophizing after the first consultation with hand surgeon suggests that clinicians change the patients' mindset during consultations and can enhance this impact to improve treatment decisions and STR. Moreover, surgically treated patients showed a greater improvement in illness perception, indicating that there is a need for a more focused strategy for changing mindset in non-invasively treated patients.

Part 3. Improve satisfaction with treatment results using data-driven tools

Based on the previous studies on factors contributing to STR, in Part 3 we developed, implemented, and evaluated two data-driven tools to improve STR. As mentioned in Part 2, various studies have highlighted the relevance and influence of mental health in musculoskeletal conditions. However, the measuring domains of mental health in clinical practice presents challenges, including increased time demands on patients and a potential lack of understanding or acceptance. Hence, in **Chapter 6**, we conducted a prospective cohort study with 19,156 patients with hand and wrist conditions to develop the Ultra-short Screener for Mental Health (pain catastrophizing, psychological distress,

and illness perception) with a maximum of 1-2 questions per construct. The most discriminatory items were selected from the Pain Catastrophizing Scale, Patient Health Questionnaire, and Brief Illness Perception Questionnaire using machine learning. We found that the ultra-short screener, with only four items, had high construct validity and performed almost as well as the full questionnaires in explaining pain and function. The screener also had high test-retest reliability and had a median response time of only 43 seconds compared to over 4 minutes for the full questionnaires. This ready-to-use, quick, and easy screener for mental health can help to start the right conversation, manage patient expectations, and support treatment decisions.

The second tool we developed is the Patient-Specific Needs evaluation, evaluating individual needs and goals for personalized treatment and better STR in **Chapter 7**. The PSN gathers the patient's most important information need before the first visit, and evaluates the fulfillment of this information need at three months. The PSN also assesses personal treatment goals, and Personal Meaningful Gain (PMG), a novel construct evaluating the minimal improvement meaningful to the individual. We developed the PSN using an iterative mixed-methods approach, and this user-centered study included patients with hand and wrist conditions, healthcare providers, and other stakeholders. The questionnaire has five questions at baseline and two at follow-up and takes about 3 minutes to complete at baseline and less than a minute at follow-up. The PSN demonstrated good performance on relevance, understandability, completeness, and usability, and moderate to high test-retest reliability. Results also showed that patients who achieved their PMG were more satisfied with their treatment results compared to those who did not, indicating good discriminative validity. The PSN is convenient and simple to complete, making it suitable for use in daily medical care. It can be utilized to aid in decision-making, managing patient expectations, and providing personalized healthcare.

Using the patient's information need measured by the PSN, in **Chapter 8** we investigated which factors determine the fulfillment of the patients' information need and the experience with information provision for hand and wrist patients at three months. We performed a linear hierarchical regression analysis in a sample of 2712 patients for the information needs and 1884 patients for the experience with information provision. The results showed that 66% of patients rated an 8 or higher on the fulfillment of their information need. Mental health and expectations were the primary factors associated with both the fulfillment of information needs and the experience with information provision. The outcomes can be practically applied in daily medical practice by presenting positive though realistic outcomes of the treatment, subsequently improving the patient's mindset; customizing information delivery according to the patient's educational level; and tailoring the information content to meet the specific needs of patients.

To gain deeper understanding of our novel developed Personal Meaningful Gain, we evaluated the PMG in **Chapter 9** across patients with thumb base osteoarthritis (TOA), trigger finger, De Quervain's, and carpal tunnel syndrome (CTS), stratified by goal domain, and comparing surgical and nonsurgical patients. In this prospective cohort study involving 5133 patients, the PMG ranged from the lowest in TOA to the highest in CTS. Surgical patients consistently reported higher PMGs than non-invasive patients, irrespective of diagnosis. The analysis revealed twenty significant associations, with pain catastrophizing showing the strongest positive association with a higher PMG. These findings underline the complex relationships between diagnosis, treatment approach, and patient expectations, emphasizing the need for personalized healthcare strategies. Recognizing modifiable factors associated with the PMG can aid in addressing overly ambitious or unambitious improvement goals, thereby supporting shared decision-making in hand and wrist care.

Finally, **Chapter 10** addresses a gap in the use of patient-reported outcome measures in assessing treatment effectiveness and research outcomes in hand and wrist care. Existing metrics such as the Minimally Important Change (MIC) and Patient Acceptable Symptom State (PASS) offer generic thresholds for determining clinically meaningful improvement but may not be well-suited for individual patients. We aimed to evaluate whether attaining the PMG better predicts STR compared to MIC or PASS, focusing on commonly used PROMs (sub)scores in hand or wrist conditions. In a prospective cohort study using eleven diverse patient samples, we assessed the positive predictive value of the PMG against various PROM (sub)scores. The results demonstrated that the PMG consistently outperformed both PASS and MIC in identifying satisfied patients across analyses, emphasizing its superior ability as a clinical outcome threshold for treatment success. Thus, implementing the PMG in clinical practice and research can enhance the accuracy of treatment evaluations, aligning with the principles of patient-centered and value-based healthcare.

In conclusion, this thesis advocates for a personalized healthcare approach, offering valuable insights and tools to enhance treatment outcomes such as STR, the patient's experience with care, and treatment decisions and evaluation. By focusing on individualized and clinically relevant outcomes, the developed tools prioritize the unique needs and goals of each patient. This approach not only acknowledges the diversity of patient needs but also empowers individuals to actively participate in their care decisions. Moreover, it helps clinicians to respond effectively. The thesis underscores the core principle that healthcare's essence lies in improving patients' lives and overall well-being, emphasizing the crucial role of individual needs, values, and goals in guiding medical decisions and interventions.

13

SAMENVATTING

Samenvatting

Het primaire doel van dit proefschrift is het bevorderen van patiëntgerichte en waardegedreven gezondheidszorg door de tevredenheid met het behandelresultaat (Satisfaction with Treatment Results, STR) van patiënten met hand- of polsaandoeningen te vergroten. Afgestemd op de principes van patiëntgerichte en waardegedreven zorg is de overkoepelende intentie om het welzijn van patiënten te verbeteren en te focussen op de cruciale rol van individuele behoeften, waarden en doelen bij het sturen van medische beslissingen en interventies.

Om dit te bereiken werden de volgende doelstellingen nagestreefd:
1. Ontwikkelen van meer begrip van tevredenheid met het behandelresultaat en gerelateerde factoren bij patiënten met hand- of polsaandoeningen
2. Verkennen van de verbinding met de mindset van de patiënt
3. Verbeteren van de tevredenheid met het behandelresultaat met behulp van datagedreven tools

Deel 1. Meten en begrijpen van de tevredenheid met het behandelresultaat

Omdat er geen gevalideerde methode was voor het meten van STR bij patiënten met hand- of polsproblemen, onderzochten we in **hoofdstuk 2** de betrouwbaarheid en validiteit van de Satisfaction with Treatment Results Questionnaire (STRQ). De STRQ werd drie maanden na aanvang van de behandeling twee keer ingevuld door 174 patiënten, en 3742 patiënten vulden de STRQ, VAS pijn en handfunctie, en de Net Promotor Score (NPS) in na drie maanden. We vonden dat de STRQ een goede tot uitstekende constructvaliditeit en hoge test-hertest-betrouwbaarheid heeft en daarom gebruikt kan worden om STR te meten.

In **hoofdstuk 3** gebruikten we vervolgens de STRQ om factoren vast te stellen die drie maanden na het begin van de behandeling in verband worden gebracht met STR. In deze cohortstudie namen we patiënten op die een carpale tunnelrelease ondergingen, een chirurgische of niet-invasieve behandeling voor duimbasisartrose, trigger finger release, beperkte fasciectomie voor contractuur van Dupuytren, of een niet-invasieve behandeling voor midcarpale laxiteit. We voerden een logistische hiërarchische regressieanalyse uit in een steekproef van 12.345 patiënten. De resultaten toonden aan dat een grotere afname van pijn tijdens belasting, een positieve ervaring met de uitleg van de voor- en nadelen van de behandeling, verbetering van de handfunctie, een positieve ervaring met advies voor thuis, betere persoonlijke controle, positieve uitkomstverwachtingen, langere verwachte ziekteduur, minder waargenomen symptomen en minder zorgen over de ziekte allemaal geassocieerd waren met een hogere STR. Onze resultaten geven aan dat om de tevredenheid van patiënten over de behandeling te vergroten, zorgverleners de ervaring van patiënten met het zorgproces kunnen verbeteren, hun

perceptie van de ziekte kunnen beïnvloeden, en hun vertrouwen in de behandeling en hun uitkomstverwachtingen kunnen vergroten.

Deel 2. Verkennen van de connectie met de mindset van de patiënt

In deel 2 verkenden we het verband tussen de mindset van de patiënt en STR, specifiek op het gebied van uitkomstverwachtingen en mentale gezondheid. Positievere uitkomstverwachtingen zijn in verscheidene onderzoeken in verband gebracht met betere behandelresultaten. De factoren die samenhangen met deze verwachtingen in de klinische praktijk worden echter niet goed begrepen, waaronder patiëntkenmerken en contextuele factoren zoals de behandeling. Het bestuderen van deze factoren zou waardevolle informatie kunnen opleveren voor het verbeteren van de uitkomstverwachtingen en, mogelijk, de behandeluitkomsten zoals STR. Daarom hebben we in **hoofdstuk 4** onderzocht welke factoren bepalend zijn voor de uitkomstverwachtingen voorafgaand aan de behandeling bij patiënten met hand- of polsaandoeningen. In deze cross-sectionele studie, waarbij we een lineaire hiërarchische regressieanalyse op verschillende niveaus uitvoerden bij 12.345 patiënten, vonden we dat een invasievere behandeling en een betere ziekteperceptie de belangrijkste factoren zijn die samenhangen met hogere uitkomstverwachtingen. De uitkomsten suggereren dat verwachtingsmanagement moet worden afgestemd op de specifieke behandeling (zoals chirurgisch versus niet-invasief) en de specifieke patiënt, inclusief hun perceptie van hun aandoening. Meer specifiek kan het nuttig zijn om zich te richten op strategieën voor verwachtingsmanagement bij niet-invasieve behandelingen, zoals handtherapie, omdat deze patiënten lagere verwachtingen hebben.

De mentale gezondheid van de patiënt speelt een cruciale rol bij het bepalen van behandelkeuzes en resultaten, zoals STR. Het verbeteren van de mentale gezondheid voorafgaand aan de behandeling zou kunnen leiden tot betere behandelbeslissingen en betere uitkomsten. Daarom evalueerden we in **hoofdstuk 5** veranderingen in de ziekteperceptie, psychische distress en pijn catastroferen van patiënten na een consult bij een handchirurg. In deze prospectieve cohortstudie vulden 276 patiënten met verschillende hand- en polsaandoeningen vragenlijsten in voor en na het eerste consult. De resultaten toonden aan dat de algemene ziekteperceptie, die verwijst naar de gedachten, overtuigingen en houdingen die iemand heeft over zijn gezondheidstoestand, verbeterde na het consult. Daarnaast verbeterden bijna alle afzonderlijke aspecten van ziekteperceptie, en deze verbetering was groter bij patiënten die een chirurgische behandeling zouden ondergaan in vergelijking met patiënten die een niet-invasieve behandeling zouden ondergaan. Verder vonden we een afname in pijn catastroferen na het consult. De verbetering van ziekteperceptie en pijn catastroferen na het eerste consult met de handchirurg suggereert dat clinici de mindset van de patiënt veranderen tijdens het consult en dit effect kunnen versterken om behandelbeslissingen en STR te verbeteren. Bovendien vertoonden chirurgische patiënten een grotere verbetering in

ziekteperceptie, wat aangeeft dat er behoefte is aan een meer gerichte strategie voor het veranderen van de mindset bij niet-invasieve patiënten.

Deel 3. Verbeteren van de tevredenheid met het behandelresultaat met datagedreven tools

Op basis van de eerdere onderzoeken naar factoren die bijdragen aan STR, hebben we in deel 3 twee datagedreven hulpmiddelen ontwikkeld, geïmplementeerd en geëvalueerd om STR te verbeteren. Zoals vermeld in deel 2, hebben verschillende onderzoeken de relevantie en invloed van geestelijke gezondheid bij aandoeningen aan het bewegingsapparaat aangetoond. Het meten van de domeinen van mentale gezondheid in de klinische praktijk brengt echter uitdagingen met zich mee, zoals meer tijd en inzet vragen van patiënten en een mogelijk gebrek aan begrip of acceptatie. Daarom hebben we in **hoofdstuk 6** een prospectieve cohortstudie uitgevoerd met 19.156 patiënten met hand- en polsaandoeningen om de Ultra Short Screener for Mental Health (pijn catastroferen, psychische distress en ziekteperceptie) te ontwikkelen met een maximum van 1-2 vragen per construct. De meest discriminerende items werden geselecteerd uit de Pain Catastrophizing Scale, Patient Health Questionnaire en Brief Illness Perception Questionnaire met behulp van machinelearning. We ontdekten dat de ultrakorte screener, met slechts vier items, een hoge constructvaliditeit had en bijna even goede prestaties opleverde als de volledige vragenlijsten in het verklaren van pijn en functie. De screener had ook een hoge test-hertest-betrouwbaarheid en had een mediane responstijd van slechts 43 seconden vergeleken met meer dan 4 minuten voor de volledige vragenlijsten. Deze gebruiksklare, snelle en eenvoudige screener voor mentale gezondheid kan helpen om het juiste gesprek te beginnen, de verwachtingen van de patiënt te managen en beslissingen over behandeling te ondersteunen.

Het tweede instrument dat we hebben ontwikkeld is de Patient Specific Needs Evaluation (PSN), waarmee individuele behoeften en doelen geëvalueerd worden voor een gepersonaliseerde behandeling en een betere STR in **hoofdstuk 7**. De PSN verzamelt de belangrijkste informatiebehoefte van de patiënt voor het eerste bezoek en evalueert de bevrediging van deze informatiebehoefte na drie maanden. De PSN beoordeelt ook persoonlijke behandeldoelen en Personal Meaningful Gain (PMG), een nieuw construct dat de minimale verbetering evalueert die zinvol is voor het individu. We hebben de PSN ontwikkeld met behulp van een iteratieve mixed-methods-aanpak en deze gebruikersgerichte studie omvatte patiënten met hand- en polsproblemen, zorgverleners en andere belanghebbenden. De vragenlijst heeft vijf vragen op baseline en twee bij follow-up en het invullen duurt ongeveer 3 minuten op baseline en minder dan een minuut bij follow-up. De PSN liet goede prestaties zien op relevantie, begrijpelijkheid, volledigheid en bruikbaarheid, en een matige tot hoge test-hertest-betrouwbaarheid. De resultaten toonden ook aan dat patiënten die hun PMG haalden, tevredener waren met hun behandelresultaten dan zij die dat niet deden, wat duidt op een goede discriminatieve validiteit. De PSN is handig en eenvoudig in te vullen, waardoor zij geschikt is voor

gebruik in de dagelijkse medische zorg. Zij kan worden gebruikt als hulpmiddel bij het nemen van beslissingen, het managen van verwachtingen van patiënten en het bieden van gepersonaliseerde gezondheidszorg.

Met behulp van de door de PSN gemeten informatiebehoefte van de patiënt hebben we in **hoofdstuk 8** onderzocht welke factoren bepalend zijn voor de bevrediging van die behoefte en de ervaring met informatievoorziening bij hand- en polspatiënten na drie maanden.

We voerden een lineaire hiërarchische regressieanalyse uit in een steekproef van 2712 patiënten voor de informatiebehoefte en 1884 patiënten voor de ervaring met informatievoorziening. De resultaten toonden aan dat 66% van de patiënten de bevrediging van hun informatiebehoefte met een 8 of hoger beoordeelde. Mentale gezondheid en verwachtingen waren de primaire factoren die samenhingen met zowel de bevrediging van de informatiebehoefte als met de ervaring met informatievoorziening. De uitkomsten kunnen praktisch worden toegepast in de dagelijkse medische praktijk door positieve maar realistische uitkomsten van de behandeling te presenteren en zo de mindset van de patiënt te verbeteren, de informatieverstrekking aan te passen aan het opleidingsniveau van de patiënt en de informatie-inhoud af te stemmen op de specifieke behoeften van patiënten.

Om meer inzicht te krijgen in onze nieuw ontwikkelde Personal Meaningful Gain evalueerden we de PMG in **hoofdstuk 9** bij patiënten met duimbasisartrose (TOA), triggerfinger, het syndroom van De Quervain en het carpaal tunnelsyndroom (CTS), gestratificeerd per doeldomein, en vergeleken we patiënten na een chirurgische of een niet-invasieve behandeling. In deze prospectieve cohortstudie met 5133 patiënten varieerde de PMG van de laagste bij TOA tot de hoogste bij CTS. Chirurgische patiënten rapporteerden consistent hogere PMG's dan niet-invasieve patiënten, ongeacht de diagnose. De analyse onthulde twintig significante associaties, waarbij pijn catastroferen de sterkste positieve associatie vertoonde met een hogere PMG. Deze bevindingen onderstrepen de complexe relaties tussen diagnose, behandelaanpak en verwachtingen van de patiënt en benadrukken de behoefte aan gepersonaliseerde gezondheidszorgstrategieën. Het herkennen van veranderbare factoren die verband houden met de PMG kan helpen bij het aanpakken van te ambitieuze of niet weinig ambitieuze verbeterdoelen, waardoor gedeelde besluitvorming in de hand- en polszorg wordt ondersteund.

Hoofdstuk 10, ten slotte, behandelt een hiaat in het gebruik van patiëntgerapporteerde uitkomstmaten bij het beoordelen van de effectiviteit van de behandeling en onderzoeksresultaten in de hand- en polszorg. Bestaande meetmethoden zoals de Minimally Important Change (MIC) en de Patient Acceptable Symptom State (PASS) bieden algemene drempels voor het bepalen van klinisch zinvolle verbetering, maar zijn mogelijk

niet goed geschikt voor individuele patiënten. We wilden evalueren of het bereiken van de PMG een betere voorspeller is van tevredenheid dan MIC of PASS, waarbij we ons richtten op veelgebruikte PROM (sub)scores bij hand- of polsaandoeningen. In een prospectief cohortonderzoek met elf verschillende patiëntpopulaties hebben we de positief voorspellende waarde van de PMG beoordeeld ten opzichte van verschillende PROM (sub)scores. De resultaten toonden aan dat de PMG consequent beter presteerde dan zowel de PASS als de MIC in het identificeren van tevreden patiënten over analyses heen, wat de superieure geschiktheid als klinische uitkomstdrempel voor behandelsucces benadrukt. Het implementeren van de PMG in klinische praktijk en onderzoek kan dus de nauwkeurigheid van behandelingsevaluaties verbeteren, in lijn met de principes van patiëntgerichte, waardegedreven gezondheidszorg.

Concluderend kan worden gesteld dat dit proefschrift pleit voor een gepersonaliseerde benadering van de gezondheidszorg en waardevolle inzichten en hulpmiddelen biedt om behandelresultaten zoals STR, de ervaring van de patiënt met de zorg en beslissingen over en evaluatie van de behandeling te verbeteren. Door zich te richten op geïndividualiseerde en klinisch relevante uitkomsten geven de ontwikkelde hulpmiddelen prioriteit aan de unieke behoeften en doelen van elke patiënt. Deze benadering erkent niet alleen de diversiteit aan behoeften van patiënten, maar stelt mensen ook in staat om actief deel te nemen aan hun zorgbeslissingen. Bovendien helpt het clinici om effectief te reageren. Dit proefschrift onderstreept het kernprincipe dat de essentie van gezondheidszorg ligt in het verbeteren van het leven en het algehele welzijn van patiënten en benadrukt de cruciale rol van individuele behoeften, waarden en doelen bij het sturen van medische beslissingen en interventies.


13

APPENDICES

List of publications

Publications in this thesis:

de Ridder WA, van Kooij YE, Vermeulen GM, Slijper HP, Selles RW, Wouters RM; and the Hand-Wrist Study Group. Test-retest Reliability and Construct Validity of the Satisfaction with Treatment Result Questionnaire in Patients with Hand and Wrist Conditions: A Prospective Study. Clin Orthop Relat Res. 2021 Sep 1;479(9):2022-2032. doi: 10.1097/CORR.0000000000001794. PMID: 34014631; PMCID: PMC8373545.

de Ridder WA, Wouters RM, Hoogendam L, Vermeulen GM, Slijper HP, Selles RW; the Hand-Wrist Study Group. Which Factors Are Associated With Satisfaction With Treatment Results in Patients With Hand and Wrist Conditions? A Large Cohort Analysis. Clin Orthop Relat Res. 2022 Jul 1;480(7):1287-1301. doi: 10.1097/CORR.0000000000002107. Epub 2022 Jan 4. PMID: 34982052; PMCID: PMC9191327.

de Ridder WA, Hoogendam L, Zeroual F, Slijper HP, Wouters RM, Vermeulen GM, Selles RW, van der Oest MJW; for the Hand-Wrist Study Group. Treatment Invasiveness and Illness Perceptions Are Strongly Associated With Outcome Expectations in Patients Treated for Hand or Wrist Conditions: A Cross-sectional Study. Clin Orthop Relat Res. 2023 May 1;481(5):994-1005. doi: 10.1097/CORR.0000000000002540. Epub 2023 Jan 24. PMID: 36727705; PMCID: PMC10097567.

Wouters RM, de Ridder WA, Slijper HP, Vermeulen GM, Hovius SER, Selles RW; Hand-Wrist Study Group; van der Oest MJW. The Ultrashort Mental Health Screening Tool Is a Valid and Reliable Measure With Added Value to Support Decision-making. Clin Orthop Relat Res. 2023 Jun 20;482(1):59–70. doi: 10.1097/CORR.0000000000002718. Epub ahead of print. PMID: 37449885; PMCID: PMC10723896.

de Ridder WA, van der Oest MJW, Slijper HP, Vermeulen GM, Hovius SER, Selles RW; Hand-Wrist Study Group; Wouters RM. Changes in illness perception, pain catastrophizing, and psychological distress following hand surgeon consultation: A prospective study. J Psychosom Res. 2023 Nov;174:111094. doi: 10.1016/j.jpsychores.2022.111094. Epub 2022 Nov 21. PMID: 37729752.

de Ridder WA, van Kooij YE, Slijper HP, Arends GR, de Roode A, MacDermid JC, Vermeulen GM, Hovius SER, Selles RW, Wouters RM; Hand-Wrist Study Group collaborators‡. Tailoring and evaluating treatment with the Patient-Specific Needs Evaluation: A Patient-Centered Approach. Plast Reconstr Surg. 2023 Dec 12. doi: 10.1097/PRS.0000000000011199. Epub ahead of print. PMID: 38085953.

de Roode A; de Ridder WA; Hoogendam L; Slijper HP; Hovius SER; Zuidam JM; Selles RW; the Hand-Wrist Study Group; Wouters RM. Which factors are independently associated with fulfilling information needs in patients treated for hand or wrist conditions? A prospective cohort study. (submitted)

Nijkamp, S; de Ridder, WA; van Kooij YE, Wouters, RM. Understanding Personal Treatment Goals towards Patient-Centered Healthcare for Patients with Hand and Wrist Conditions: a Prospective Cohort Study. (submitted)

Loos NL; de Ridder WA; Hoogendam L; MacDermid, J; Hoogeboom T; Furniss, D; The Hand-Wrist Study Group; Wouters RM. The Personal Meaningful Gain is a Better Measure of Treatment Success than the Minimal Important Change and Patient Acceptable Symptom State: a cohort study. (submitted)

Other publications:

Cohen A, Selles RW, de Ridder WA, Ter Stege MHP, Souer JS, Wouters RM; Hand–Wrist Study Group Collaborators. What Is the Impact of the COVID-19 Pandemic on Quality of Life and Other Patient-reported Outcomes? An Analysis of the Hand-Wrist Study Cohort. Clin Orthop Relat Res. 2021 Feb 1;479(2):335-345. doi: 10.1097/CORR.0000000000001514. PMID: 33044314; PMCID: PMC7899601.

Berentschot JC, de Ridder WA, Bek LM, Heijenbrok-Kal MH, Braunstahl GJ, Remerie SC, Stuip Y, Ribbers GM, Aerts JGJV, Ista E, Hellemons ME, van den Berg-Emons RJG; CO-FLOW Collaboration Group. Patients' evaluation of aftercare following hospitalization for COVID-19: satisfaction and unmet needs. Respir Res. 2024 Mar 29;25(1):145. doi: 10.1186/s12931-024-02748-8. PMID: 38553722; PMCID: PMC10981299.

van Kooij YE, Ter Stege MHP, de Ridder WA, Hoogendam L, Hovius SER, MacDermid JC, Selles RW; Hand-Wrist Study Group; Wouters RM. The Validity and Responsiveness of the Patient-Specific Functional Scale in Patients With First Carpometacarpal Osteoarthritis. J Hand Surg Am. 2024 Jun 25:S0363-5023(24)00197-7. doi: 10.1016/j.jhsa.2024.04.006. Epub ahead of print. PMID: 38934999.

Arends GR, Loos NL, van Kooij YE, Tabeau K, de Ridder WA, Selles RW, Veltkamp J; Outcome-Based Healthcare Research Group; Wouters RM. What are the perspectives of patients with hand and wrist conditions, chronic pain, and patients recovering from stroke on the use of patient and outcome information in everyday care? A Mixed-Methods study. Qual Life Res. 2024 Jun 5. doi: 10.1007/s11136-024-03685-1. Epub ahead of print. Erratum in: Qual Life Res. 2024 Jul 27. doi: 10.1007/s11136-024-03724-x. PMID: 38839681.

Arends GR, van Kooij YE; Stern BZ; Hovius SER; Wouters RM; de Ridder WA. Which factors make patients with hand and wrist conditions dissatisfied with their treatment results despite achieving their personal goals? A qualitative study. (submitted)

Authors and affiliations

Ruud W. Selles
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.

Harm P. Slijper
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Xpert Clinics Hand and Wrist Care, the Netherlands.

Guus M. Vermeulen
- Xpert Clinics Hand and Wrist Care, the Netherlands.

Robbert M. Wouters
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.

Yara E. van Kooij
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Center for Hand Therapy, Xpert Handtherapie, The Netherlands.

Lisa Hoogendam
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Xpert Clinics Hand and Wrist Care, the Netherlands.

Steven E.R. Hovius
- Dept. of Plastic Surgery, Radboud UMC, Nijmegen, the Netherlands.
- Former head of Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Xpert Clinics Hand and Wrist Care, the Netherlands.

Mark J.W. van der Oest
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Xpert Clinics Hand and Wrist Care, the Netherlands.

A. (Leanne) de Roode
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Fysio- & Oefentherapie Rhenen-Elst (Utrecht), The Netherlands

Sharon Nijkamp
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Physical Therapy Sciences Program in Clinical health Sciences, University Medical Center Utrecht, Utrecht University, the Netherlands

Nina L. Loos
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.

G.R. (Danée) Arends
- Dept. of Rehabilitation Medicine, Erasmus MC, Rotterdam, the Netherlands.
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.
- Center for Hand Therapy, Xpert Handtherapie, The Netherlands.

Fadoua Zeroual
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.

Thomas J. Hoogeboom
- Radboud university medical center, IQ Health science department, Nijmegen, Netherlands.

Dominic Furniss
- Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom.

Joy C. MacDermid
- School of Physical Therapy, Western University, London, ON, Canada.
- School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada.
- The Hand and Upper Limb Centre, St Joseph's Health Centre, London, ON, Canada.

J. Michiel Zuidam
- Dept. of Plastic and Reconstructive, and Hand surgery, Erasmus MC, Rotterdam, the Netherlands.

PhD portfolio

PhD Portfolio

Summary of PhD training and teaching

Name PhD student: Willemijn A. de Ridder	PhD period: Aug 2019 – March 2023
Erasmus MC Department: Rehabilitation Medicine & Plastic, Reconstructive and Hand Surgery	Promotor(s): Prof. Dr. Ruud W. Selles
Research School: NIHES	Supervisor: Dr. Harm P. Slijper, Dr. Guus M. Vermeulen

1. PhD training

	Year	Workload (ECTS)
General courses		
Introduction to Medical Writing	2021	2
LLS Scientific Integrity	2021	0.3
LLS Leadership and Teamwork	2021	0.3
LLS Intervision	2020	0.4
Research Integrity	2021	0.3
Using R for Statistics in Medical	2020	1.4
Personal Leadership and Communication	2021	0.7
Public Speaking, Speech Republic	2022	0.7
BKO (Basis Kwalificatie Onderwijs)	2022	5.1
BKE (Basis Kwalificatie Examinering)	2022	1.8
Research BROK ('Basiscursus Regelgeving Klinisch Onderzoek')	2023	1.5
Specific courses (e.g. Research school, Medical Training)		
Study Design	2019	4.7
Biostatistical Methods I: Basic Principles	2019	5.7
Biostatistical Methods II: Classical Regression Models	2019	4.3
Principles of Research in Medicine and Epidemiology	2019	0.7
Methods of Public Health Research	2019	0.7
Health Economics	2019	0.7
Introduction to Global Public Health	2019	0.7
The Practice of Epidemiologic Analysis	2019	0.7
Fundamentals of Medical Decision Making	2019	0.7
Advances in Clinical Epidemiology	2020	0.7
Repeated Measurements in Clinical Studies	2020	1.7
Review of Mathematics and Introduction to Statistics	2020	1
Causal Inference	2021	2.0
Causal Mediation Analysis	2021	3.7
Clinical Translation of Epidemiology	2021	1.4
Introduction to Bayesian Methods in Clinical Research	2021	1.4
Clinical Epidemiology	2021	3.7
Implementation Science	2021	1.4
Principles in Causal Inference	2022	1.4
Seminars and workshops		
Research meetings Hand Wrist Study Group, Erasmus MC	2019-2023	4.1

	Year	Workload (ECTS)
Oral scientific presentations	2019	0.5
Dupuytren Vereniging:		
Post-operative treatment of Dupuytren's disease	2021	0.5
Handtherapie met hart en data, Xpert Handtherapie, Utrecht:		
Satisfaction with treatment results: fact and fiction		
Federation of European Societies for Surgery of the Hand (FESSH) annual meeting, Rotterdam (online):	2021	1
Which factors are associated with satisfaction with treatment result in patients with hand or wrist conditions? A Large Cohort Analysis		
Test-Retest Reliability and Construct Validity of the Satisfaction with Treatment Result Questionnaire in Patients with Hand and Wrist Conditions: a Prospective Study		
Annual International Musculoskeletal Mental and Social Health Consortium (I-MESH) symposium (online)	2022	0.5
Treatment invasiveness and illness perceptions drive outcome expectations in patients treated for hand or wrist conditions. A cohort study		
Federation of European Societies for Surgery of the Hand (FESSH) annual meeting, London:	2022	1.5
Best paper award session: Treatment invasiveness and illness perceptions drive outcome expectations in patients treated for hand or wrist conditions. A cohort study		
Changes in illness perception, pain catastrophizing, and psychological distress following hand surgeon consultation: a prospective cohort study		
The development of the patient-specific needs questionnaire: a patient-reported tool evaluating individual patient needs, treatment goals and improvement goals		
IFSSH/IFSHT combined congress 2023, Rimini, Italy:	2023	0.7
Boosting outcome expectations: why, when and how (invited lecture)		
(Inter)national conferences attendance		
FESSH Copenhagen, Denmark	2018	1.1
IFSSH/IFSHT Berlin, Germany	2019	1.1
Data science in a day, Hand Wrist Study Group Rotterdam	2020	0.3
FESSH Rotterdam, Netherlands (online)	2021	1.1
FESSH London, UK	2022	1.1
Nationale Wetenschapscommunicatiedag, Den Haag	2022	0.3
IFSSH/IFSHT, Rimini, Italy	2023	1.1

2. Teaching

	Year	Workload (ECTS)
Lecturing and supervising		
Systematic review – minor students	2021-2022	1.5
Bachelor 3rd year anatomy (practicum)	2021-2023	1.5
Master student's anatomy (practicum)	2021-2023	1.5
Minor bachelor students, anatomy of the hand and forearm, injuries of the upper extremity, hand muscle strength (lectures)	2021-2023	1.5
Master's student, Mijke Vreeken	2019-2020	1.4
Master's student, Leanne de Roode	2021-2022	1.4
Master's student, Sharon Nijkerk	2022-2023	1.4
Master's student, Danée Arends	2022-2023	1.4
Researcher, Fadoua Zeroual	2021-2022	1.4
Other		
Review FWO Vlaanderen grant applications	2021	0.7
Review article for the Journal of Pain	2021	0.2
Handtherapie met hart en data, Xpert Handtherapie, Utrecht, participating as a data scientist in a hackathon	2022	0.3
Science communication:		
Developing a format for science communication, sharing our research on socials, website and other	2021, 2022	3.3
Organizing annual Symposium Handtherapie met Hart en Data	2021, 2022	1.5
Total		88 ECTS

About the author

On December 6th, 1984, Willemijn de Ridder came into the world in a quaint white house in Harich, Friesland, just a day after Sinterklaas, as her mother jokingly claimed she “held her back” to avoid a birthday coinciding with the festive celebration. Willemijn grew up in a small town between the farmers and the woods, where her family of seven constituted 1,75% of the total population of the village. Growing up in the middle of nowhere gave her a lasting love for walking, nature, animals, and traveling the path less traveled.

In her youth, Willemijn dreamed of becoming an Olympic gymnast, yet soon realized her temperament clashed with the rigid discipline demanded by coaches in the era. Nonetheless, she performed as a (acrobatic) gymnast and dancer till after her student years. Alongside her athletic pursuits, Willemijn nurtured other career aspirations, including a desire to specialize in sports physical therapy or pursue a path in journalism and writing. Her passion for literature led her to win awards for her poems and self-publish her high school thesis, “Weg, weg, weg,” an anthology of short stories and poems.

Willemijn’s wide-ranging interests guided her to study Languages and Cultural Studies at Radboud University Nijmegen. She enthusiastically took part in the university’s Honours Program and did a program focused on public speaking and negotiation, which sparked her interest in international politics. This led her to spend a semester studying political science at the University of Vienna.

Accepted to the master’s program in Conflicts, Territories and Identities at the Radboud University, Willemijn acquired an internship at the Ministry of Foreign Affairs, where she delved into the realm of diplomacy. She combined her interest in literature and politics in her master’s thesis on forgiveness in South Africa and graduated cum laude. Still yearning for more knowledge, she further studied theology and philosophy on a Thomas More grant, developing a newfound fascination with metaphysics.

Rekindling her desire to contribute to people’s health, Willemijn started studying physical therapy part-time while balancing jobs as a writer, translator, gymnast teacher, and yoga instructor. She then started working as a hand therapist at Xpert Clinics, a specialized clinic for hand surgery and therapy, swiftly becoming the manager of three hand therapy teams.

Motivated by her innate curiosity and ambition, Willemijn embarked on a PhD journey at the Hand-Wrist Study Group of Erasmus MC, University Medical Center Rotterdam. During

this time, she and her husband Robbert celebrated the arrival of sons Izo and Felix into their lives. While balancing the demands of both academia and parenthood, Willemijn managed to earn a master's degree in Clinical Epidemiology. Their life became a (most of the time) joyful blend of research, family, and the occasional chaos that comes with raising children.

Presently, Willemijn serves as a policy officer for the program [Ont]Regel de Zorg at the Ministry of Health, Welfare and Sports. Dedicated to reducing the administrative burden on healthcare workers, her mission is to make and maintain data-driven, personalized, and high-quality care in the Netherlands accessible, fair, and affordable for everyone.

Dankwoord

Hoe vaak je ook ging
Even vaak kwam je aan
Ieder spoor blijft bestaan
Waar je was legt je vast

Geen stap die je zet
Wordt ooit over gedaan
Je bent hier nog niet weg
Of je komt er vandaan

Ingmar Heytze, straatgedicht op de Biltstraat in Utrecht

Dit boek is bijna uit. Ik kan hier nu wat schrijven over een hell of a ride, of mijn promotietraject vergelijken met het beklimmen van een berg, met het schilderen van een schilderij, of het baren van een kind. Maar dat doe ik niet. Het was het schrijven van een boek. En daarmee vervul ik (deels) een lang gekoesterde wens. Ik ben trots op het resultaat, waarmee ik daadwerkelijk iets bijdraag aan betere zorg. Daarnaast ben ik trots op mezelf en mijn omgeving, dat ik het volbracht heb. Naast het krijgen van en zorgen voor twee fantastische kinderen, het behalen van nog een master, een verbouwing en verhuizing (twee keer) en de COVID-19 pandemie. Ik ben een tevreden mens.

Ik had dit nooit kunnen doen zonder de hulp van een heleboel geweldige mensen, die ik nu eindelijk mag bedanken.

Prof. Dr. Ruud Selles, beste Ruud, tijdens het schrijven van dit dankwoord ging ik erover nadenken wat het zo prettig maakt om met jou samen te werken. Hier komt ie: 1. Je maakt heerlijke pizza. 2. Je bent een erg prettige leidinggevende, die bovenal meedenkt. Je hebt me altijd het gevoel gegeven dat je echt wilde dat ik iets leerde en me ontwikkelde als onderzoeker en werknemer. 3. Je bent een steengoede onderzoeker en schrijver, die elk onderzoek en artikel zóveel beter maakt. Je hebt me altijd de ruimte gegeven om zelf na te denken over het hoe, wat en waarom. 4. Je bent een heel gezellige collega en vriend: ik hoop nog regelmatig met jou en Marjolijn (en natuurlijk onze andere ZEER gewaardeerde collega) ergens een hapje te eten, (whisky) te drinken en te praten over de grote dingen des levens, zoals voetbal, politiek en emancipatie.

Dr. Harm Slijper, beste Harm, zonder jou was dit boekje er niet geweest. Jij hebt ervoor gezorgd dat dit onderzoek gedaan mocht worden, en dat ik dan ook nog eens degene was die het onderzoek mocht gaan uitvoeren. Je hebt elk artikel en ieder onderzoeksidee kritisch bekeken, bevraagd, omgedraaid, weggegooid, weer opgepakt, bespuugd, opgepoetst, tentoongesteld. Dit was een fascinerende werkwijze waar standaard iets

moois uit voortkwam. Naast allerlei inhoudelijke zaken heb ik ook van je geleerd hoe ik vrij dominante mensen kan onderbreken om zelf iets te kunnen zeggen. Een skill voor het leven dus.

Dr. Guus Vermeulen, beste Guus, jij was het chirurgische mes in mijn onderzoek. Je hebt mijn papers korter, bondiger, klinisch en pragmatisch gemaakt. Ook je praktische tips voor het schrijven van de algemene stukken en voor het opvoeden van twee kids kwamen altijd van pas.

Leden van de leescommissie, prof. dr. Maaike Kleinsmann, prof. dr. Jan van Busschbach en prof. dr. Cindy Veenhof: veel dank voor het lezen en beoordelen van dit proefschrift. Het is een eer om de bevindingen uit dit proefschrift met jullie te mogen delen en bediscussiëren.

Heel veel dank aan **Xpert Clinics**, voor het faciliteren van mijn ambities en voor het mogen uitvoeren van dit onderzoek. Speciale dank aan **Rob van Huis**. Ik weet nog dat ik van jou moest kiezen: management of onderzoek. Ik koos management, maar ik herken ook een kans als die zich voordoet. Voor iemand die altijd een boek heeft willen schrijven had ik bij deze keuze geen beslisondersteuning nodig om me te helpen. En ik weet dat jij weet dat dit de juiste keuze was. Daarom dank voor je support bij mijn ontwikkeling bij Xpert Clinics: van therapeut tot manager tot onderzoeker.

Alle chirurgen van Xpert Clinics hand- en polszorg: dank voor de leuke dagen. Ik heb het als leerzaam en gezellig ervaren om samen met jullie controles te doen. Diezelfde controles waren ook aanleiding voor mij om promotieonderzoek te gaan doen. De beslissingen die we nemen is voor een groot deel gebaseerd op gevoel; mooi om dat te kunnen rationaliseren. Met name **Xander en Sebastiaan**: jullie hebben me de kneepjes van het vak geleerd. Ook dank aan de **revalidatieartsen** van Xpert Clinics, met name **Kirsten** voor het oproepen van vragen nadat alle vragen beantwoord leken te zijn.

Em. prof. dr. Steven Hovius, beste Steven, als handtherapeut werkte ik al erg graag met je samen. Voor mij waren de controles samen altijd een combinatie tussen les krijgen en gezelligheid. Als onderzoeker was dat niet anders. Dank voor je steun, waardering en interesse. De groetjes thuis, hè.

Mijn oud-collega handtherapeuten van Xpert Handtherapie: dank jullie wel! Niet alleen voor het gezamenlijk wegdrinken van onze hersencellen maar ook voor de verhitte discussies over handen en witte broeken. Een heel aantal van jullie beschouw ik als mijn vrienden, en spreek ik nog steeds. Een speciale high five voor mijn **oude team Leiden**: met jullie wil ik elke dag wel samenwerken.

Alle leden van de Hand-Wrist Study Group: ik kan jullie niet genoeg bedanken voor de goede opmerkingen, de interessante onderwerpen en de ruimte om wel of niet iets te zeggen. Ik hoop jullie in de toekomst nog te zien, voorlopig op congressen of promoties, en anders op de pizza-avonden tot we allemaal grijs of kaal zijn.

Dank **dr. Mark van der Oest**, ik was altijd onder de indruk van hoe je kennis, kunde en snelheid combineerde met plezier. Ik vond het prettig om met je samen te werken omdat ik er altijd iets van leerde en natuurlijk omdat je een extreem snelle jongen bent.

Promovendi van de 15e en 16e: we hebben sowieso te weinig geborreld en (duur) geluncht. Maar de keren dat we het wel hebben gedaan waren onbetaalbaar. Waarvoor dank!

Dank **prof. dr. Irene Mathijssen en prof. dr. Gerard Ribbers** voor de mogelijkheid om aan jullie afdelingen te mogen promoveren. Ook dank aan de **collegae van revalidatiegeneeskunde en plastische chirurgie** voor de het fijne contact op de afdelingen.

Many thanks to **all co-authors** in this thesis for your input and for improving our work.

Dank aan mijn gemotiveerde **studiegenoten van NIHES**, en in het bijzonder **Paul Werthmann** für seine begeisterte Unterstützung. Without the walks with you it would have been a lot less fun. Together we discovered Rotterdam, a city we have taken to our hearts.

Ook dank aan mijn oud-collega's van het **Kennisinstituut van de Federatie Medisch Specialisten**: in de korte periode dat ik bij jullie werkte hebben jullie mijn onderzoek meer gewicht gegeven.

Collega's van het ministerie van VWS, in het bijzonder **team [Ont]Regel de Zorg en mijn buddy Renske**: jullie hebben mijn perspectief op de gezondheidszorg verbreed. Ook hebben jullie me met open arme ontvangen en langzaam aan de gebruiken van een ministerie laten wennen; een heel andere wereld en tegelijkertijd vergelijkbaar met de complexe wereld van het onderzoek. Ik kijk er naar uit de boel te blijven [ont]regelen samen.

Lieve **vrienden**, ik heb jullie veel te weinig gezien! Maar nu het boekje eindelijk af is ontstaan er natuurlijk weer zeeën van tijd voor samen eten, drinken, discussies over sport, politiek, emancipatie en anders leuks. Speciale dank aan mijn bestie **Simone** voor het jarenlange luisterende oor. Ook wil ik de **buurtjes van de Tureluur** (en de Tortel, aangezien daar mijn favoriete combinatie van oud-collega's/vrienden/buren woont) bedanken: jullie hebben me afgeleid, precies op de momenten dat ik het nodig had. En al gaan we verhuizen, zie jullie woensdag om 16u bij de picknicktafel. Ook wil ik mijn **sportmatties** bedanken: voor het samen in en uit het zweet werken, de borrels, uitjes, tips

voor activiteiten in de buurt... En natuurlijk voor het uitwisselen van de tops en vooral flops van de week. Gelukkig is alles met Sally, Roxanne en de balletjes van de koningin beter.

Lieve Lisa Hoogendam, we zijn samen aan het promotietraject begonnen en haalden samen onze masterdiploma's op. Onze totaal verschillende karakters en interesses vulden elkaar mooi aan en zorgden ervoor dat we elkaar op meerdere momenten verder hielpen. Ik had het ook graag samen met je afgesloten en je naast me gehad tijdens mijn verdediging maar je hebt iets veel mooiers te doen (/te verzorgen). Dat maken we dan wel weer goed in de binnenspeeltuin.

Lieve Yara van Kooij, mijn paranimf! Robbert noemde jou zijn partner in crime, dus dat zal ik niet herhalen. Maar ik zie jou wel als mijn kompaan tijdens dit hele traject. Jouw rustige, weloverwogen manier van denken en doen is inspirerend. Als ik met jou heb gesproken hoef ik die dag niet meer te mediteren. Dat heeft zowel mijn werk als mijn leven verbeterd. Ik ben blij dat ik je mag blijven bellen als het dreigt mis te gaan met mijn mindset.

Lieve Hilde Koster, toen ik je vroeg of je mijn paranimf wilde zijn, vroeg je: "Denk je dat ik dat kan?" Ja, er is niets wat jij niet kan dus ik kan met een gerust hart flauwvallen tijdens mijn verdediging. Jij bent met stip mijn meest uitgesproken vriendin en je bent een van de mensen met de meeste invloed op mijn meningen over zeer uiteenlopende zaken (politiek en emancipatie inderdaad). Dus paranimf, vorm je mening en sta klaar om voor me in te vallen.

Lieve Cees, José, Ruud, Christine, Tom, Judith en Roos: dank voor de (bij tijden terecht voorzichtige) interesse in mijn onderzoek. En voor het weten wanneer je er niet naar moest vragen. En voor het oppassen op de kinderen, zodat ik even bij kon komen. Ik prijs me gelukkig met jullie om me heen.

Lieve papa en mama, als antropoloog gaven jullie het juiste voorbeeld: nieuwsgierig blijven. Binnen ons gezin was er altijd ruimte om vragen te stellen, om interesses en talenten te ontplooien, om iets uit te proberen. En heel veel vrijheid. Niet zo gek dus, dat promotieonderzoek. Wel gek dat het in deze richting is, als je m'n oude schoolrapporten bekijkt. Van alfa via gamma naar bèta. Ik wist niet dat ik het in me had. Maar ik heb wel altijd de ruimte gevoeld om vastbesloten van richting te veranderen, en het vertrouwen dat elke richting de juiste is. En toch wel leuk om de tweede dr. De Ridder te zijn.

Ida, Jeroen, Bregt, Niels, Rieneke, Anna en Guido: wat fijn dat jullie me (onbewust) altijd hebben uitgedaagd om tot het gaatje te gaan. Het lot van het vierde kind misschien. Never give up, never surrender. Nooit verliezen. Vervelende eigenschap bij het spelen van spelletjes, maar een kwaliteit voor het afronden van dit boekje. **Merel, Benthe, Feija, Thijmen, Ingmar en Sanna**: wat fijn dat jullie er zijn!

Lieve Izo en Felix, nu het boekje af is heb ik nog meer tijd om voor te lezen en handstanden te doen. Dank voor jullie vrolijkheid, (poepiedanpa)grapjes, liedjes, (snot) kusjes en knuffels. Jullie hebben mijn vermogen tot relativeren vergroot en mijn leven verrijkt. Wat ben ik dol op jullie!

Lieve, lieve Robbert, ik wist al dat we een goed team zijn, maar de werkelijkheid heeft mijn verwachtingen overtroffen (zie je: hoge verwachtingen leiden tot betere uitkomsten). We zetten natuurlijk wel wat op het spel door zo nauw samen te gaan werken. Vanwege de COVID-19 pandemie werd het nog nauwer dan we hadden kunnen vermoeden. Ik denk dat ik oprecht kan zeggen dat het onze relatie alleen maar heeft versterkt. Ik weet niet hoe je het gedaan hebt: koffie gemaakt wanneer dat echt nog het enige was dat kon helpen, me ervan weerhouden mijn laptop uit het raam te smijten, geholpen wanneer nodig, maar ook gelaten wanneer noodzakelijk. Aangemoedigd en afgeremd. Vastgehouden en losgelaten. Liefde. En nu is deze periode echt voorbij en zijn we geen collega's meer, maar wel allebei doctor. En die titel is net als onze liefde voor altijd ;) *Geen stap die je zet wordt ooit overgedaan.* Op naar de volgende stap in ons fantastische leven, op weg naar nieuwe avonturen, want samen kunnen wij alles! #RWWR4ever