Further knowledge

HOW TO USE PATIENT-REPORTED OUTCOMES MEASURES AND OTHER CLINICAL MEASUREMENTS IN CLINICAL REPORTS

Clinical outcomes in hand surgery can be measured and reported in many ways, including those performed under controlled conditions by clinicians, such as grip strength or finger joint angles, and assessments of function that involve completing specific simulated daily tasks. Patient-reported outcome measures (PROMs) are another group of outcome measures, which are defined as reports obtained directly from the patient, and typically involve questionnaires. PROMs can assess hand function, health status, and patient satisfaction.

The use of PROMs is rising in scientific articles in hand surgery. However, PROMs are not necessarily a panacea. Here, we will aim to rationalise the choice of outcome measures using our current understanding of PROMs in hand surgery.

What PROMs reflect

The World Health Organisation International Classification of Functioning, Disability and Health (WHO ICF) defines: (1) Body structures and functions: "impairment" results from their disruption or pathology. (2) Activities: "limitation" arises when they cannot be undertaken. (3) Participation: "restriction" occurs when this is affected (http://www.who.int/classifications/icf/en/). It can be appreciated how different hand surgery outcome measures fit into this hierarchy. Abnormal finger joint angles represent "*impairment*" of the anatomical joint structure and function, poor Sollerman hand function test performance amounts to "*limitation*" in activities, and high scores on work and social activity questions in the Disabilities of the Arm, Shoulder and Hand (DASH) PROM describe "*restriction*" of participation. In this hierarchy, PROMs can assess "impairment" (e.g. questions about stiffness or paraesthesia), "limitation" (e.g. questions about holding cutlery), or "restriction" (e.g. questions about not attending work or social activities).

Harnessing this WHO ICF-based classification, a series of ordered priorities can be considered to guide the choice of outcome measures when designing research, writing a report, or critiquing an article.

First, consider the question to be answered and the audience to be targeted

Different scientific questions in clinical studies require different outcome measures. Taking randomised controlled trials (RCTs) as examples, a study of the efficacy of a new drug for rheumatoid arthritis is towards the explanatory end of a spectrum ranging from "explanatory" to "pragmatic" studies. In such a trial, finger joint motion might be measured, which is a clinician-derived measurement of "impairment". Alternatively, measuring patient satisfaction via a PROM would be challenging, as it would be difficult to separate the biological impact of the drug from things like the placebo effect.

In contrast, in a pragmatic study of a hand therapy regimen, a PROM assessing hand function would be more appropriate, as this would give a comprehensive assessment of the impact that the intervention has had on the patient. This is because the net benefit that a patient experiences from hand therapy can stem from gaining education about activities that can be performed safely, and having increased confidence about returning to work, in addition to an improvement in the finger joint range of motion. In this way, measuring finger joint range of motion alone would not "tell the whole story".

Explanatory and pragmatic trials of the same condition may have different primary outcome measures. Table 1 summarises two RCTs of the management of Dupuytren's disease that illustrate such a difference.

Table 1: Summary of two trials of the management of Dupuytren's disease

Both RCTs study collagenase in Dupuytren's disease, yet they have different designs and use different outcome measures to answer different questions. The choice of primary outcome measures reflects these differences and aligns with the aim of the study in each. If the outcome measures were switched between the two trials, the data generated in each would be less meaningful.

Some research assumes that measures of "impairment of structure or function" are effective surrogates for "limitation in activity" or "restriction from participation". However, this cannot be assumed to be the case. For example, in Dupuytren's disease, finger joint angles do not correlate with hand function, as assessed with PROMs, as shown in several studies (Engstrand et al., 2009, Jerosch-Herold et al., 2011). Ensuring adequate patient and participant involvement when designing the study may reduce the risk of making inappropriate assumptions about how different measures are related to each other, and which of them matter to patients.

As alluded to in the aforementioned examples, different research questions will target particular audiences. The CORD-1 trial, an explanatory study of collagenase versus placebo, might demonstrate efficacy to a regulator, who is aiming to establish whether collagenase should be licensed. However, its findings may be less useful to a clinician and patient attempting to make a treatment decision between collagenase and open surgery. In contrast, the ongoing DISC trial is pragmatically comparing collagenase to fasciectomy and will provide clinically-relevant data that is more likely to impact the clinician-patient consultation. The value of the outcome measures in the two studies relates to the needs of their respective audiences. When designing or appraising studies, initial attention ought to be directed to the specific question and the target audience. This will inform what kind of outcome measure is required.

Second, evaluate the quality of PROMs and meaningfulness of changes and differences

When choosing a PROM for a given study or report, the quality of the available options should be considered. Currently, discussion of outcome measures (and PROMs in particular) often involves describing them as “validated” or “not validated”. This is an oversimplification for two major reasons. First, the clinical situation for which a PROM has been validated may differ from the intended use. Second, validity is not a single entity that can be passed or failed, just as describing a car as “good” or “bad” is an oversimplification, as a car can be judged on power, comfort, economy, reliability, and other domains. Both of these points can be considered in the following hypothetical examples:

1. A hand function PROM that includes several items about pain may have been validated in a painful condition, like thumb basal joint arthritis. However, it may not perform well in a painless condition, like radial nerve motor palsy. To extrapolate its validity from studies of thumb basal joint arthritis when designing a trial of radial nerve palsy treatment is inappropriate.

2. A PROM that has been shown to have good reliability (only one of the elements of validity), is not completely “validated”. Other aspects of its psychometric properties, such as whether the items it comprises measure the same underlying trait or construct (e.g. hand function), have not been studied. Thus, the PROM has been only partially validated.

There are tools for judging the quality of outcome measures. These are typically structured as checklists that can be used to “tick off” the different elements of validity for a PROM, building a complete understanding of its psychometric properties (Dobbs et al., 2018, Terwee et al., 2007). These tools were primarily designed to assess PROMs, but they can be used to assess other outcome measures as well. For example, extension deficit as an outcome measure (as used in the CORD-1 trial) may be relevant as it assesses the efficacy of collagenase at correcting contractures, but it is not comprehensive for assessing the global

benefit or harm to patients. Thus, extension deficit might tick a checklist box for relevance, but not one for comprehensiveness. Of the two quality assessment systems cited, one provides a stepwise approach to selecting a suitable outcome measure (Dobbs et al., 2018), and has similarities to what is outlined in Figure 1.

Figure 1: Suggested steps in selecting a suitable outcome measure

Having selected an outcome measure, it is important to ensure the study is powered to detect a clinically meaningful change for an individual, or difference between groups.

Third, establish the roles of primary and secondary outcomes

Setting the roles of PROMs and other measures to be used is the next step in study design. A study that uses a PROM as the primary outcome measure may include other measures, for example grip strength, as secondary outcome measures. How the PROM and other measures will be used to support the interpretation of the results should be considered.

Currently available PROMs

A range of PROMs exist that can be used in hand surgery, and broadly include (1) PROMs that are generic and can be used across all conditions (e.g. the EuroQol 5D, which assesses health status), (2) site-specific PROMs that are applicable to an anatomical area (e.g. the DASH PROM already mentioned, which can be used across upper limb conditions), and (3) disease-specific PROMs that are applicable to a single condition (e.g. the Boston Carpal Tunnel Questionnaire). By following a stepwise approach to outcome measure selection like that proposed here, the most suitable type of PROM can be chosen.

The use of PROMs in many studies is proper and necessary, and our understanding of how to use them is evolving. However, confirmation of the quality of existing PROMs is needed. It should be appreciated that all outcome measures have limitations, which need to be

systematically assessed and accounted for in the study design and appraisal processes, and this field remains an important area for future hand surgery research.

Funding Statement

JNR is an Honorary Specialty Registrar in Plastic Surgery and an NIHR Postdoctoral Fellow. The views expressed are the authors' own, and are not necessarily those of the NIHR, NHS, or Department of Health.

References:

Dobbs TD, Hughes S, Mowbray N, Hutchings HA, Whitaker IS. How to decide which patient-reported outcome measure to use? A practical guide for plastic surgeons. J Plast Reconstr Aesthet Surg. 2018, 71: 957-66.

Engstrand C, Boren L, Liedberg GM. Evaluation of activity limitation and digital extension in Dupuytren's contracture three months after fasciectomy and hand therapy interventions. J Hand Ther. 2009, 22: 21-6.

Jerosch-Herold C, Shepstone L, Chojnowski A, Larson D. Severity of contracture and self-reported disability in patients with Dupuytren's contracture referred for surgery. J Hand Ther. 2011, 24: 6-10.

Terwee CB, Bot SDM, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007, 60: 34-42.

Jeremy N Rodrigues[1*], Carlos Neblett[2]

[1]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Oxford, UK.

[2]University Hospital of The West Indies, Kingston, Jamaica

*corresponding author: Jeremy.rodrigues@ndorms.ox.ac.uk