Seminars in Arthritis and Rheumatism 43 (2014) 721–729

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Reliability of joint count assessment in rheumatoid arthritis: A systematic literature review

Peter P. Cheung, MBBS, FRACP, FAMS[a,b,d,*], Laure Gossec, MD, PhD[c], Anselm Mak, MD, FRCP, FAMS[a,b], Lyn March, MBBS, FRACP, MSc (Epidem & Biostat), FAFPHM, PhD[d]

[a] Division of Rheumatology, National University Health System, Singapore
[b] Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[c] Department of Rheumatology, UPMC Univ Paris 06, GRC-UPMC 08 (EEMOIS), AP-HP, Pitié-Salpêtrière Hospital, Paris, France
[d] Department of Rheumatology, University of Sydney, Institute of Bone and Joint Research, Royal North Shore Hospital, St Leonards, Australia

ARTICLE INFO

Keywords:
Synovitis
Rheumatoid arthritis
Health care professional and patient joint counts
Reliability
Reproducibility
Clinical examination

ABSTRACT

Background: Joint counts are central to the assessment of rheumatoid arthritis (RA) but reliability is an issue.
Objectives: To evaluate the reliability and agreement of joint counts (intra-observer and inter-observer) by health care professionals (physicians, nurses, and metrologists) and patients in RA, and the impact of training and standardization on joint count reliability through a systematic literature review.
Methods: Articles reporting joint count reliability or agreement in RA in PubMed, EMBase, and the Cochrane library between 1960 and 2012 were selected. Data were extracted regarding tender joint counts (TJCs) and swollen joint counts (SJCs) derived by physicians, metrologists, or patients for intra-observer and inter-observer reliability. In addition, methods and effects of training or standardization were extracted. Statistics expressing reliability such as intraclass correlation coefficients (ICCs) were extracted. Data analysis was primarily descriptive due to high heterogeneity.
Results: Twenty-eight studies on health care professionals (HCP) and 20 studies on patients were included. Intra-observer reliability for TJCs and SJCs was good for HCPs and patients (range of ICC: 0.49–0.98). Inter-observer reliability between HCPs for TJCs was higher than for SJCs (range of ICC: 0.64–0.88 vs. 0.29–0.98). Patient inter-observer reliability with HCPs as comparators was better for TJCs (range of ICC: 0.31–0.91) compared to SJCs (0.16–0.64). Nine studies (7 with HCPs and 2 with patients) evaluated consensus or training, with improvement in reliability of TJCs but conflicting evidence for SJCs.
Conclusion: Intra- and inter-observer reliability was high for TJCs for HCPs and patients: among all groups, reliability was better for TJCs than SJCs. Inter-observer reliability of SJCs was poorer for patients than HCPs. Data were inconclusive regarding the potential for training to improve SJC reliability. Overall, the results support further evaluation for patient-reported joint counts as an outcome measure.

Introduction

Rheumatoid arthritis (RA) is an inflammatory disorder characterized by synovitis, a process of joint inflammation that leads to joint destruction, and ultimately to physical disability [1]. Early detection and treatment of synovitis has been shown to reduce radiographic progression and has been part of the overarching principles of "treating-to-target," with the ultimate objective of achieving and remaining in remission [2]. Hence, it is important that physicians are able to reliably detect clinical synovitis.

Joint count assessment by physicians through swollen and tender joints is considered the most conventional way of detecting clinical synovitis [3], and its importance in disease activity assessment is supported by its inclusion in core data sets of disease activity indices such as the Disease Activity Score in 28 joints (DAS28) [4] and the American College of Rheumatology (ACR) response criteria [5] used in clinical trials, research, and clinical practice. In addition, it is part of the Outcomes in Measures for Rheumatoid Arthritis in Clinical Trials (OMERACT) RA Core Set [5].

As a measurement outcome, joint counts can measure disease activity in RA [5], are sensitive to change [6], predictive of radiographic progression [5,7], and correlate with other surrogate markers of disease activity [8]. Although physicians traditionally perform joint counts, nurses and other health professionals (metrologists) have started to participate, either as part of extended roles or for clinical research [9]. In addition, the association of patient-reported swollen

* Correspondence to: Division of Rheumatology, University Medicine Cluster, National University Health System, 1E Kent Ridge Rd, Tower Block Level 10, 119228, Singapore.
E-mail address: peter_cheung@nuhs.edu.sg (P.P. Cheung).

0049-0172/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.semarthrit.2013.11.003

and tender joints when compared with joint counts derived by physicians or metrologists has been evaluated [10]. This increased interest is due to the evidence that regular assessment of disease activity is important for patients to achieve and remain in remission [2]. Specifically, detection of residual swollen joints is clinically important for patients in clinical remission (DAS28 < 2.6), as this contributes significantly to continual radiographic progression [7]. Hence, monitoring especially between clinic visits with joint counts will potentially optimize achievement of "treating-to-target" [11].

However, the reliability by both the same person performing the assessment (intra-observer) and between different assessors (inter-observer) is debatable, particularly the inter-observer reliability [12,13]. It is unclear how reliable physician- or patient-reported joint counts are and whether reliability of patient-reported joint counts is comparable to that between health care professionals (i.e., physician or metrologist). Training through didactic teaching or standardization through group consensus following set guidelines has been used to improve reliability in various clinical outcome measures in rheumatology [14,15]. It is uncertain whether this can improve reliability in joint count assessments.

The objective of this systematic review was to evaluate the reliability and agreement of joint counts (intra-observer and inter-observer) by physician, metrologists, or patients and to assess whether training or standardization may improve joint count reliability and agreement in these studies.

Methods

Literature search strategy

A systematic search was performed in PubMed, EMBASE, and Cochrane databases up to September 2012. Inclusion criteria were articles reporting reliability or agreement of joint counts or articular indices in RA derived by physicians, metrologists, or patients. The search was limited to humans and publications in English and French. The following exploded medical subject heading (MeSH) terms were used in PubMed: "Arthritis, Rheumatoid" AND either "Joints" OR "severity of illness index" OR "physicians" OR "patients" AND either "reproducibility of results" OR "observer variation" OR "reliability" OR "analysis of variance."

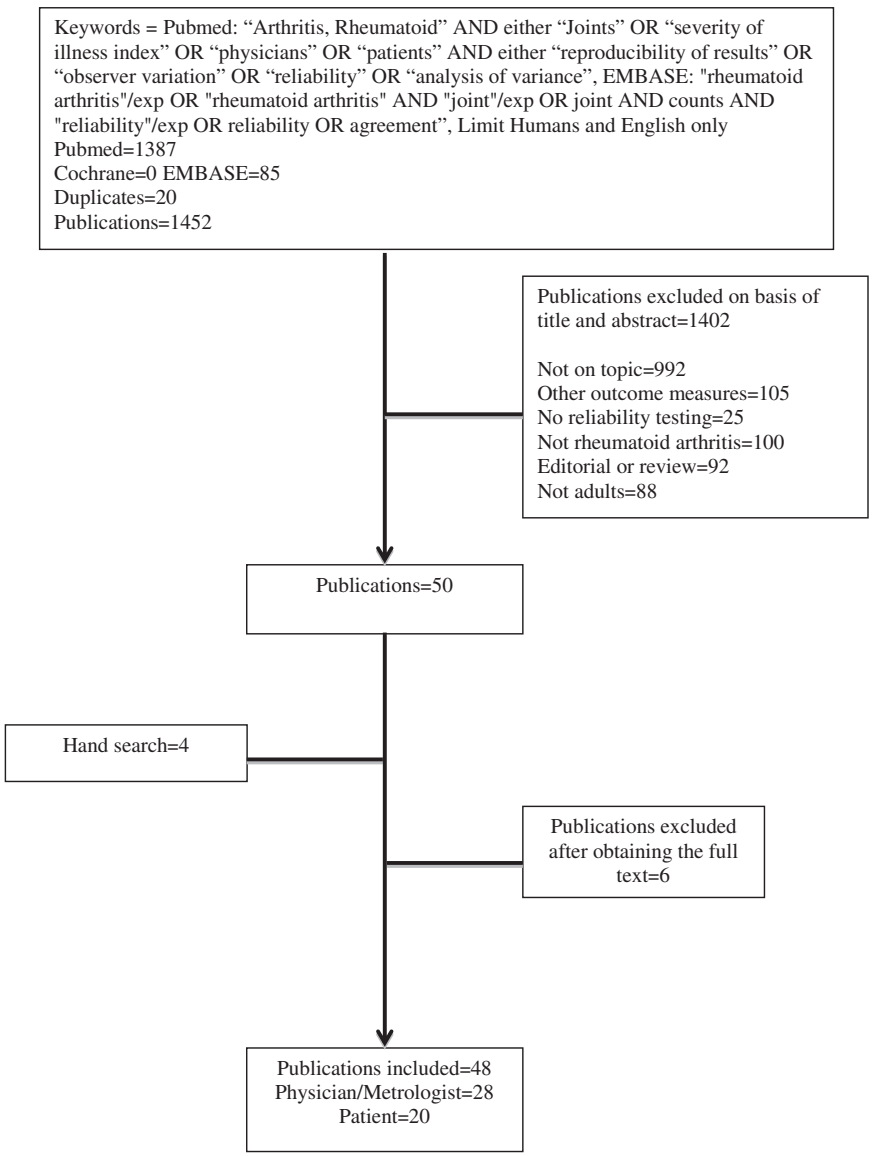

Fig. 1. Literature search strategy.