



Universiteit  
Leiden  
The Netherlands

## Use of multidimensional composite scores in rheumatology: parsimony versus subtlety

Landewe, R.B.M.; Heijde, D. van der

### Citation

Landewe, R. B. M., & Heijde, D. van der. (2021). Use of multidimensional composite scores in rheumatology: parsimony versus subtlety. *Annals Of The Rheumatic Diseases*, 80(3), 280-285. doi:10.1136/annrheumdis-2020-216999



Version: Publisher's Version

License: [Creative Commons CC BY-NC 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/3196081>

**Note:** To cite this publication please use the final published version (if applicable).

# Use of multidimensional composite scores in rheumatology: parsimony versus subtlety

Robert B M Landewé <sup>1,2</sup>, Désirée van der Heijde <sup>3,4</sup>

**Handling editor** Josef S Smolen

<sup>1</sup>Amsterdam Rheumatology Center, Amsterdam University Medical Center, Amsterdam, The Netherlands

<sup>2</sup>Rheumatology, Zuyderland Medical Center, Heerlen, The Netherlands

<sup>3</sup>Rheumatology, Leiden University Medical Center, Leiden, The Netherlands

<sup>4</sup>Rheumatology, Diakonhjemmet Hospital, Norway

## Correspondence to

Professor Robert B M Landewé, Amsterdam Rheumatology Center, AMC, 1105 AZ Amsterdam, The Netherlands; landewe@rlandewe.nl

Received 24 August 2020

Revised 20 October 2020

Accepted 20 October 2020

Published Online First

3 November 2020

## ABSTRACT

Rheumatic and musculoskeletal diseases (RMDs) form a diverse group of diseases. Proper disease assessment is pivotal, for instance to make treatment choices and for optimising outcome in general. RMDs are multidimensional diseases, entrenching many, sometimes very different aspects. Composite outcome measures ('composites') have become very popular to assess RMDs, because of their claim to catch all relevant dimensions of the disease into one convenient measure. In this article we discuss dimensionality of RMDs in the context of the most popular conceptual framework of RMDs, being an inflammatory process leading to some sort of damage over time. We will argue that multidimensionality not only refers to heterogeneity in disease manifestations, but also to heterogeneity in possible outcomes. Unlike most unidimensional measures, multidimensional composites may include several disease manifestations as well as several outcome dimensions into one index. We will discuss fundamental problems of multidimensional composites in light of modern strategies such as treat-to-target and personalised medicine.

Finally, we will disentangle the use of multidimensional composites in clinical trials versus their use in clinical practice, and propose simple solutions in order to overcome problems of multidimensionality and to avoid harm to our patients due to overtreatment.

## INTRODUCTION

Rheumatic and musculoskeletal diseases (RMDs) form a diverse group of diseases with different pathogeneses.<sup>1</sup> The prevalence of several RMDs is relatively high (>1%) and most RMDs are chronic diseases. Treatment options have recently expanded for some RMDs but are still sparse or absent for others. Treat-to-target strategies have become popular.<sup>2</sup> Proper disease assessment is pivotal for physicians to make choices about treatment start, intensification or tapering, and optimising outcome in general. The wish to choose the best and most (time) efficient instrument is understandable. Efficiency here implies capturing as much as possible by one 'simple' measure. This is why composite outcome measures have become so popular. Here we will investigate their rationale further, discuss the concept of dimensionality and warn against some misuses.

## COMPOSITE INDICES

As a reflection of the wish to bring some order in a profusion of single outcome measures, composite indices have found their place in rheumatology. A 'composite' combines several measures into one quantifiable index, which is a rather generic

principle,<sup>3,4</sup> that is visualised in [figure 1](#). In theory, a composite index is better than the sum of its parts, but this assumption is hard to prove and sometimes not met.<sup>5</sup> If one single measure does not satisfactorily describe what is going on in most patients, if not in all, one could use multiple single measures that all reflect the same process to some extent. But multiple measures create multiple problems. If separate measures give diverse signals, which one then reflects the truth best? What potentially important aspect of a disease will be missed by making exclusive choices? What if among five single measures for improvement, three suggest improved disease activity and two do not? For a well-designed 'composite', developers must have thought critically about these problems. They must have achieved consensus on questions like what exactly to address, which variables to include and exclude (prioritisation), how do these variables correlate, how should variables be weighted, among others. It is not easy to design a 'good composite'. It is even more difficult, if not political, to obtain common support for a new index, so that it will be implemented.

Advocates of composites tend to believe that several instruments put together smartly give a better picture of the situation than only one instrument would do. Disease activity in rheumatoid arthritis (RA) can be measured by a plethora of different single measures. More pain (eg, on a visual analogue scale) may point to more active diseases, as does a higher swollen joint count, an increased C reactive protein (CRP) level and the patient's global impression of the disease. But not all patients with RA with active joints report similarly high levels of pain, while some with many swollen joints may have a normal CRP or no pain at all, and patients often rate their disease as being more active than their physicians do. The merging of different perspectives of the same domain into one index may sometimes add clarity and uniformity, and help clinical research and practice move forward as we have seen in the last three decades, but there are certainly also problems.

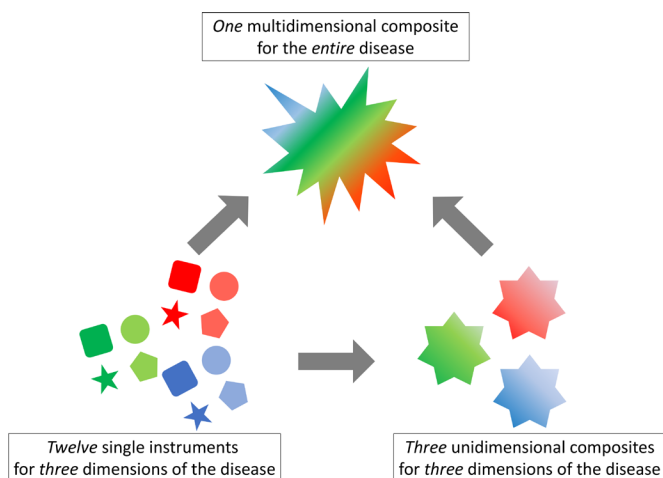
## RMDs ARE MULTIDIMENSIONAL DISEASES

Patients with RMDs usually have musculoskeletal symptoms and sometimes extra-musculoskeletal manifestations. These latter can be organ specific or more diffuse, and may involve several internal organs. RMDs have many faces; they are multifaceted or multidimensional. Some RMDs, such as systemic lupus erythematosus (SLE) and psoriatic arthritis (PsA), are classical examples of multidimensionality. Phenotypically, they may express a multitude of manifestations, but infrequently in



© Author(s) (or their employer(s)) 2021. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Landewé RBM, van der Heijde D. *Ann Rheum Dis* 2021;**80**:280–285.



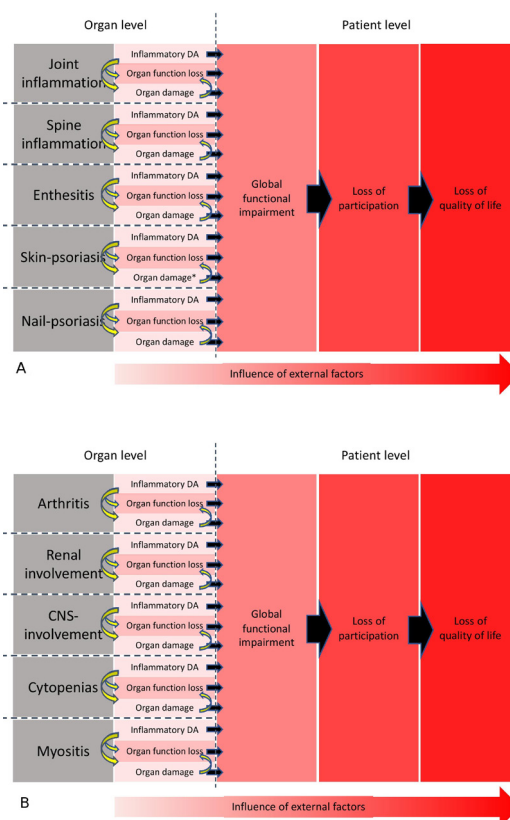
**Figure 1** From single instruments to multidimensional composites; a reductionistic approach. The figure visualises how several single instruments, that are grouped to reflect different dimensions of the same disease (red–blue–green), may be aggregated into unidimensional composite indices or scores, by a data-driven process of prioritisation, exclusion and weighting (dark-grey arrows). Single instruments and unidimensional composite scores can further be aggregated into one multidimensional composite index or score by a similar process of prioritisation, exclusion and weighting. Different colours reflect different dimensions. Different tones reflect different perspectives (*bright*: patient's perspective, *dull*: physician's perspective). Increasing irregularity of the symbols reflects increasing versatility of measurements, increasing shades of colour reflects increased dimensionality.

the same patient at the same time. Two patients diagnosed with the same disease may present very differently; there is marked *between-patient variability*, which has implications for properly assessing these patients. Good disease measures can discern this level of heterogeneity in all its possible extremes.

Most RMDs are chronic and rarely stable. They fluctuate in symptom intensity, naturally or under the influence of treatment; there is marked *within-patient variability*. Rheumatologists need to pick up these fluctuations in order to adjust treatment. Good disease measures can pick up these fluctuations reliably.

Multidimensionality does not only exist at the level of disease presentation, but expectedly also at the level of disease outcomes. Disease activity is an immediate outcome of many inflammatory RMDs. On top of that, patients with RMDs face a gradual accumulation of chronic and irreversible *consequences* of their disease (activity) over time. Examples are, among others, progressive joint destruction, increasing functional impairment or atherosclerosis. These consequences can be seen as dimensions too, but in a perpendicular orientation. **Figure 2** provides a schematic representation of multidimensionality of RMDs in the opinion of the authors: phenotypical dimensions along the y-axis and dimensions of outcome along the x-axis.

Who wants to describe and understand the breadth of outcomes of an RMD must capture both the disease *process* and the consequences of that process, but we rather tend to simplify things. Categorising outcomes into analysable dichotomies, such as *responses* or *events*, which is often done in randomised trials, is an impoverishment, since most of the natural variability gets lost. The outcome of an RMD is usually not an event, such as a myocardial infarction or death, but rather a quantification of an ongoing disease process characterised by fluctuations that say a lot about the disease and the patient. Dichotomising outcomes



**Figure 2** Multidimensionality of rheumatic musculoskeletal diseases (RMDs). The complexity of two multidimensional RMDs, psoriatic arthritis (PsA) (A) and systemic lupus erythematosus (SLE) (B). Note that, while disease manifestations of PsA and SLE are very different, both diseases share a large degree of multidimensionality. Phenotypical manifestations of the diseases are placed along the y-axis (in grey), and consequences of the disease are depicted along the x-axis (in red). The figure suggests two types of hierarchy, one with vertical orientation and one with horizontal orientation. The vertical hierarchy refers to organ (systems) and reflects the associations that exist between organ inflammation, organ damage and organ function loss (yellow arrows). The horizontal hierarchy reflects a natural gradient from 'organ' (left) to 'organism' (right). The larger coloured arrow outside the figure represents the multitude of external factors that may have an impact on the outcomes along the x-axis. The gradient of increasing 'redness' reflects increasing distance to the primary process underlying the RMD, and increasing sensibility to external, not necessarily disease-related factors. Examples are comorbidities, sensitivity to central sensitisation, personality traits, usage of coping mechanisms and illness perceptions, among others. The summing up of disease manifestations and outcomes in the figure is not intended to be complete, and any suggestion pointing to a hierarchy in severity of manifestations is unintentional. \*Inflammatory psoriasis lesions will usually heal without damage (scars), but may lead to itching, scratching and scarring (indirect damage). CNS, central nervous system; DA, disease activity.

into digestible binomial parcels provides statistical convenience and comprehension, but does not give sufficient credit to the complexity of RMDs. Still, we often do this, for reasons of simplicity, and obviously for buying time in a busy clinic.

#### FRAMEWORK: PROCESS AND DAMAGE

The conceptual framework underlying many of our RMDs is that immunological disturbances cause inflammation. The process of inflammation gives measurable clinical signs (eg, joint swelling) and symptoms (eg, pain, stiffness) instantaneously, and

irreversible structural organ destruction (damage) after a while. Many of our RMDs are not necessarily inflammatory RMDs. However, the conceptual framework, with inflammation as the *process* and damage as the *consequence*, has become so axiomatic that we have extended this label to all RMDs, even when inflammation as a cause is less clear. The degenerative disease osteoarthritis and the pain syndrome with the anachronistic name fibrositis owe their suffix *-itis* to this type of generalisation rather than to clear evidence that inflammation is key.

The inflammation-damage framework has been instrumental in the development of rheumatology as it stands today. First, the framework provided the insight that in order to avoid irreversible damage inflammation should be suppressed, an insight that has made way for successful drug development. Second, the framework stood model for the hypothesis of ‘window-of-opportunity’; it appreciated the importance of ‘time elapsed’ which led to the paradigm of ‘starting an intervention sooner rather than later’. *Time-is-joint*. Third, the framework has shaped the field of outcome assessment of RMDs. Both process and damage (note: damage in its broadest sense) can now be measured appropriately by a wealth of instruments. As in every cause–effect relationship, a proper interpretation of the temporal association between process and damage is essential. Simply stated, disease activity comes first and damage follows after some time. The interpretation of disease activity and damage at the same time, while ignoring the time elapsed as in a cause–effect relationship, conveys different signals. Disease activity happens *now*, damage is a remnant of a process in the *past*. As we will see later, some composites neglect the importance of time elapsed, and mix things up.

### MULTIDIMENSIONALITY AND OUTCOME ASSESSMENT

The word dimension can be used to describe one *aspect* out of a spectrum. Myositis can be considered one dimension of the disease SLE, and cytopenia (haematological manifestation) another one. Skin psoriasis is one dimension of PsA, nail psoriasis another one and arthritis a musculoskeletal one. The word *dimension* can also be used to describe one *outcome* out of a spectrum of possible outcomes. Joint damage can be considered one dimension out of the spectrum of possible outcomes of PsA, and reduced quality of life another one. Not all patients with PsA and joint damage, however, will perceive and report reduced quality of life over time, or will lose their job due to the disease. External factors will largely determine to what extent proximal outcome variables measured at the organ level (eg, joint inflammation, joint damage) will ultimately impact quality of life and well-being (figure 2). The distinction between several dimensions is arbitrary and based on expert convention.

Rheumatologists have an irresistible desire to behold a multidimensional RMD as a whole, and to treat the patient with this RMD in its entirety; rheumatologists are ‘lumpers’ by soul and advertise the holistic view. No wonder that they have developed multidimensional composite measures that account for the whole patient, covering all aspects of the disease and its outcomes into one measure (see figure 1). Because of their presumed user-friendliness (*one size fits all*), multidimensional composites enjoy significant popularity for use in trials and increasing attractiveness in clinical practice. Clinically relevant trade-offs allow categorisation into *disease states* and *response states*, which get an intuitive meaning among clinicians over time. One example of such a multidimensional composite is the Psoriatic Arthritis Disease Activity Score (PASDAS).<sup>6</sup> Minimal Disease Activity (MDA), a threshold, is conceptually

a composite measure developed to be used as a treatment target in the same disease.<sup>7</sup> One of the many examples of multidimensional composites for SLE is the Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI), that captures many dimensions of the disease in one index.<sup>8</sup> Dichotomous derivatives of SELENA-SLEDAI include definitions for mild, moderate and severe flares. Those several multidimensional indices (such as Systemic Lupus Activity Measure, Lupus Activity Index, British Isles Lupus Assessment Group Index and the European Consensus Lupus Activity Measure, among others, (reviewed by Mikdashi and Nived)) that have seen the light over the years exemplify that multidimensionality of a disease does not necessarily add to consensus on how to measure the disease best.<sup>9</sup>

### UNIDIMENSIONAL OUTCOME MEASURES

Unidimensional composites differ from multidimensional ones in that they cover only one aspect (‘dimension’) of the disease (figure 1). Many are in use for measuring the state or change of disease activity. Sometimes, it is not immediately clear whether a composite measure is unidimensional or multidimensional. A closer look at its history may give some resolution. A hallmark dimension of RMDs is joint inflammation (arthritis). The Disease Activity Score (DAS) was developed in 1990 for assessing disease activity in RA and clearly focused on arthritis.<sup>10</sup> It was obvious that one single measure (eg, a swollen joint count or an acute-phase reactant) would not suffice to appropriately describe disease activity in every patient with RA. The DAS has been set up as a composite index combining several measures covering the same dimension. DAS in its origin was a unidimensional index with a focus on (the immediate sequels of) arthritis. This does not imply, however, that once unidimensional means always unidimensional, as the following example may clarify. In the last decades, the *Gestalt* of RA has changed, due to earlier recognition, more effective treatment and better management.<sup>11</sup> As a consequence, among others, average inflammatory burden is assumed to be lower now than it was in the past. However, recent studies have suggested that the gradual decrease in swollen joint count and acute-phase reactants over time did not go hand in hand with less patient-reported pain, less joint tenderness and more well-being.<sup>12 13</sup> Part of this discrepancy is currently attributed to the existence of neuropathic pain mechanisms or central sensitisation.<sup>14</sup> Pain due to central sensitisation falls outside the conceptual inflammation-damage framework, although one may provocatively argue that central pain sensitisation is a long-term consequence of inflammation, and thus damage. Anyway, neuropathic pain constitutes a different dimension of RA than pain that accompanies inflammation. Indeed, this type of pain is rather insensitive to anti-inflammatory drug treatment, and does not correlate with CRP and swollen joint count. That means: DAS, once a unidimensional composite for disease activity in patients with active RA who had to start treatment,<sup>10</sup> may have gained dimensions over time, when used to monitor patients with RA in remission or in low disease activity. Exactly the same reasoning pertains to the DAS-lookalikes Simple Disease Activity Index<sup>15</sup> and Clinical Disease Activity.<sup>16</sup> That this may have implications for daily clinical care has been demonstrated by us and others recently in studies comparing DAS28 and the fully patient-reported index RAPID3 in all day practice.<sup>12 13 17</sup> Apparently, the context in which a measure has been developed versus the context in which it is used is relevant for a proper understanding of the measure’s performance. Obviously, similar issues may happen with other measures in other diseases.



**Table 1** Summary of advantages and disadvantages of single measures, unidimensional composites and multidimensional composites

	Advantages ('pros')	Disadvantages ('cons')
Single measures	Cost little time per measure, are easy to use and clearly interpretable (eg, 'VAS pain: 3/10')	Provide limited information per measure
	Many can be measured independently in the same patient	Prioritisation may lead to interpretational problems and <i>cherry picking</i>
	Recognise individual patients with extreme values and not-so-average problems	
Unidimensional composites	Provide an unambiguous picture of one dimension (eg, 'SDAI showing LDA')	'Dimension creep' may happen over time (eg, neuropathic pain impacts SDAI scores)
		May mix up process and damage variables (eg, HAQ score in the ACR response)
	Provide more statistical power by eliminating variability (smoothing)	Dampen the influence of extreme values that are not recognised as such anymore
	Provide feasible and consensual benchmarks for treat-to-target strategies	Benchmarks may have a different meaning in different stages of the disease
Multidimensional composites	Presume a holistic and unambiguous picture of the patient (eg, 'the patient is doing well')	Aggregate measures and composites based on statistical considerations ('lumping')
		May suffer from the index within the index fallacy which jeopardises feasibility (eg, HAQ score and PASI score in MDA)
		May mix up process and damage variables (eg, HAQ score in the MDA)
	Provide more statistical power by less variability (smoothing)	Assume different dimensions change always in the same direction (eg, skin and joints in PsA)
		Lose statistical (discriminatory) power if components of the composite or index are not correlated (eg, SLEDAI)
	Provide benchmarks for treat-to-target strategies (eg, 'the patient is not yet in MDA')	Benchmarks may have a different meaning in different patients, do not give resolution about which dimension to treat, and their use may lead to overtreatment
		Benchmarks may have a different meaning in different stages of the disease

Advantages and disadvantages are paired as much as possible.

ACR, American College of Rheumatology; HAQ, Health Assessment Questionnaire; LDA, low disease activity; MDA, minimal disease activity; PASI, Psoriasis Activity and Severity Index; PsA, psoriatic arthritis; SDAI, Simple Disease Activity Index; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; VAS, visual analogue scale.

## DIMENSIONS, CORE DOMAINS AND INSTRUMENTS

Although there is certainly overlap, it is important to distinguish multidimensionality of RMDs from core domains, such as the ones operationalised by the Outcome Measures in Rheumatology Clinical Trials (OMERACT) organisation.<sup>18</sup> OMERACT has always aimed at clinical trials and has approached outcome assessment in rheumatology from the perspective of best (ie, feasible, discriminative and truthful) measurement. OMERACT makes a distinction between 'what to measure' (the core domains) and 'how to measure' (the best instruments). 'What to measure' refers to a conceptual framework that is accepted among all stakeholders as the 'truth'. Core domains can be very diverse and are supposed to represent so called core areas, such as death, life impact, resource use and economic impact, pathophysiological manifestations and adverse events. Certain aspects of the RMD (dimensions) may not pop up in OMERACT core-domain sets, for example because they cannot be measured well, or have a too low prevalence.

In summary, multidimensionality is a feature of RMDs. It requires a conceptual framework to explain the disease phenotypically, its pathogenetic causes and its longitudinal consequences. Whether these dimensions should be assessed or not in trials is the focus of OMERACT. OMERACT core sets increase the comparability across studies, which is pivotal, but do not aim at providing completeness.

Thus far, OMERACT has not taken an explicit stand with regard to the use of composite indices, but has allowed some composites as preferable instruments for assessing some of their core domains. Many of these composites, however, had been developed long before they were 'pulled through the OMERACT filter' and they have often been accepted under stakeholder pressure, since 'they are important to patients' or 'they work satisfactorily in the context of clinical trials'. Important limitations were either not realised or ignored. We will discuss a few.

## IGNORING THE NATURAL ORDER OF CAUSE AND CONSEQUENCE

Composite indices should respect the natural order of cause and consequence, as argued above. Some indices used in rheumatology violate this principle. The American College of Rheumatology (ACR) response measure ACR20, endorsed by many regulatory bodies and OMERACT, was designed as a response measure for RA disease activity, but includes a measure of functional ability (the Health Assessment Questionnaire (HAQ)).<sup>19</sup> HAQ measures functional impairment as a *consequence* of RA disease activity, not disease activity itself. The PASDAS, a measure for disease activity, also includes the HAQ. Who looks at the content of the HAQ realises that all kinds of conditions, not only RA-related or PsA-related disease activity, may influence HAQ score. It is true that HAQ score correlates reasonably well with direct measures of disease activity in patients who have active disease, but we do not know how this works out in patients who are inactive or have only mild disease activity, nor in those that actually have pain without clinical signs of inflammation. Studies have shown that HAQ incorporates an irreversible component that proportionally increases over time,<sup>20</sup> which implies that an HAQ score in a patient with early active RA does not have the same meaning as the same HAQ score in a patient with quiescent but advanced disease. As such, the HAQ as a part of a response index should be considered methodologically inappropriate.

## THE 'INDEX WITHIN AN INDEX' FALLACY

The aforementioned ACR20 response measure, an index, includes the HAQ, an index itself. The aforementioned MDA includes the HAQ and the Psoriasis Area Severity Index for skin involvement.<sup>7</sup> Indices tend to dampen the influence of extreme values and reduce variability. This helps increasing the signal-to-noise ratio, which is a statistical advantage, but goes at the cost of subtlety necessary to properly assess individual patients with non-classical presentations. This smoothing process is reinforced by dichotomising clinical

outcomes, as in ACR20 response (present or absent) and MDA state (present or absent), among others. While indices, with components that are indices themselves entrenched, may still work for groups of patients in randomised trials, they are essentially useless in describing and monitoring individual patients, unless these patients belong to the *typically averaged*. ‘Useless’ becomes ‘potentially dangerous’ if dichotomised multidimensional composites, such as MDA, are used as targets for intensifying treatment. Too many other factors than inflammatory disease activity alone may have impact on whether or not a patient meets a preset threshold. Overtreatment is the logical consequence of threshold medicine, when inappropriate measures to ascertain the threshold are used.

The dangers of threshold medicine also pertain to unidimensional indices, but to a lesser extent, and, besides, these unidimensional composites less often suffer from ‘the index within an index’ fallacy.

### A PATIENT IS NOT THE SUM OF HIS DIMENSIONS

A patient with SLE who has active myositis does not necessarily have other manifestations of SLE. Change in the activity of myositis is not necessarily related to change in—for example—leucocyte count, skin rash or arthritis. Along similar lines, the relationship between psoriasis skin activity and musculoskeletal symptoms of PsA is modest at best.<sup>21</sup> That some drugs used for PsA may improve both skin and joints, and have indications for both, does not mean that skin and joint can be expected to always change in similar directions. Multidimensionality does not imply that separate dimensions, present at the same time, change at similar speed or in similar direction. Still, a multidimensional index pretends to allow a unidimensional (ie, linear) interpretation. Patients with scores above the threshold are ‘not good’; only those with scores below the threshold are ‘good’. Two patients with PsA, however, may have similar levels of PASDAS but very different manifestations and burden of disease (impact). Their response to treatment may also markedly differ. In groups of patients in randomised trials, this may work out to some extent, as long as experimental therapies have unidirectional positive effects on several dimensions. But in diseases like SLE, systemic sclerosis or primary Sjögren’s, multidimensionality of outcome and response measures may obscure clinically relevant heterogeneity among patients. One of the potential explanations for failed trials with drugs that experienced physicians perceive as efficacious in patients with SLE, indeed pertains to this kind of heterogeneity that is inherent to the composite outcome measures. Multidimensionality can jeopardise sensitivity to change and discrimination.

### IMPLICATIONS

Advocates of multidimensional composites will argue that these validated indices have sufficiently proven their value, but what does validation mean? Indeed, many of these composites have worked reasonably well in randomised trials, in that they can distinguish between groups on active treatment versus those on placebo. However, that is low-hanging fruit. As argued above, these indices tend to eliminate the outlier effects by statistical smoothing, resulting in better signal-to-noise ratios, more statistical power and better p values. Problems may arise, however, if results of randomised controlled trials are to be generalised to common clinical practice. It is uncertain whether the statistically significant result obtained with a multidimensional composite in a trial keeps up in patients with the same disease but a somewhat different, not-so-average clinical presentation or course. Problems may also arise if the trial has exploited a benchmark, for example, in the context of treat-to-target, and the strategy that includes the measure and the benchmark is implemented one-to-one in clinical practice. Patients that would benefit from intensification of treatment, since they have measurable

residual inflammatory activity in one dimension, may not be picked up as such. More likely, though, benchmarks may falsely dictate further intensification of treatment in patients who have responded well to treatment in several dimensions, but fail to do so in a variable that reflects the consequence of a process (eg, function impairment) rather than the process itself (eg, inflammation). Further intensification of anti-inflammatory treatment may not necessarily improve these patients’ lives meaningfully. Consequent overtreatment will make medicine unnecessarily costly and risky.<sup>22</sup> Benchmark medicine with suboptimal multidimensional instruments is pointless.

### HOW CAN WE DO BETTER?

Composite indices have their value in clinical trials. A better signal-to-noise ratio adds to statistical power and limits the numbers of patients needed in the trial. However, composite indices should give credit to the complexity of the disease, not by trying to lump all dimensions into one index, but rather to respect the *time elapsed* between cause and consequence and avoid mixing both up. The versatility of an RMD is better valued by reporting different dimensions differently, as for example propagated by Schoels *et al* for PsA.<sup>23</sup> Multiple unidimensional indices likely are better tools for purpose than one single multidimensional index, but when the breadth of outcomes is relevant, it may be even better to describe the separate components of the composite as secondary outcomes, in conjunction with the unidimensional index itself.

When application of composites in clinical practice is in question, multidimensional composites lose their value, since interpretational mistakes are too easily made, and patients may fall victim to benchmark medicine. Unidimensional indices likely perform better, but also bear the risk of mixing up different perspectives into one index. The classic example is the patient with RA with high DAS but low inflammation. When decisions about treatment start, intensification or tapering are to be made, physicians should realise the rationale of (anti-inflammatory) treatment: to reduce the process of inflammation in order to avoid long-term consequences of the disease. This means, the measure to base such decisions on should linearly reflect the presence of (objective) inflammation. Swollen joint counts and acute-phase reactants, or a physician with real experience in detecting inflammation clinically, may do a better job here than composites. A summary of advantages and disadvantages of composite measures, unidimensional and multidimensional, versus single measures is provided in [table 1](#).

Parsimony in outcome assessment can unintentionally lead to loss of subtlety and harm rather than benefit patients in clinical practice.

**Contributors** Both authors have discussed the work, written the manuscript and approved the final version.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not required.

**Provenance and peer review** Commissioned; externally peer reviewed.

### ORCID iDs

Robert B M Landewé <http://orcid.org/0000-0002-0577-6620>

Désirée van der Heijde <http://orcid.org/0000-0002-5781-158X>

### REFERENCES

- van der Heijde D, Daikh DI, Betteridge N, *et al*. Common language description of the term rheumatic and musculoskeletal diseases (RMDs) for use in communication with the lay public, healthcare providers and other stakeholders endorsed by the European League against rheumatism (EULAR) and the American College of rheumatology (ACR). *Ann Rheum Dis* 2018;7:829–32.

- 2 Smolen JS, Aletaha D, Bijlsma JWJ, *et al.* Treating rheumatoid arthritis to target: recommendations of an international Task force. *Ann Rheum Dis* 2010;69:631–7.
- 3 Greco S, Ishizaka A, Tasiou M, *et al.* On the methodological framework of composite indices: a review of the issues of weighting, aggregation, and robustness. *Soc Indic Res* 2019;141:61–94.
- 4 Organisation for economic cooperation and development (OECD). Handbook on constructing composite indicators; methodology and user guide. Available: [www.oecd.org/sdd/42495745.pdf](http://www.oecd.org/sdd/42495745.pdf)
- 5 Lukas C, Landewé R, Sieper J, *et al.* Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;68:18–24.
- 6 Helliwell PS, FitzGerald O, Fransen J, *et al.* The development of candidate composite disease activity and Responder indices for psoriatic arthritis (grace project). *Ann Rheum Dis* 2013;72:986–91.
- 7 Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. *Ann Rheum Dis* 2010;69:48–53.
- 8 Castrejón I, Tani C, Jolly M, *et al.* Indices to assess patients with systemic lupus erythematosus in clinical trials, long-term observational studies, and clinical care. *Clin Exp Rheumatol* 2014;32:S-85–95.
- 9 Mikdashi J, Nived O. Measuring disease activity in adults with systemic lupus erythematosus: the challenges of administrative burden and responsiveness to patient concerns in clinical research. *Arthritis Res Ther* 2015;17:183.
- 10 van der Heijde DM, van 't Hof MA, van Riel PL, *et al.* Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;49:916–20.
- 11 Sparks JA, Arthritis R, Med AI. Rheumatoid arthritis. *Ann Intern Med* 2019;170:ITC1.
- 12 Ferreira RJO, Duarte C, Ndosi M, *et al.* Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. *Arthritis Care Res* 2018;70:369–78.
- 13 Ferreira RJO, Dougados M, Kirwan JR, *et al.* Drivers of patient global assessment in patients with rheumatoid arthritis who are close to remission: an analysis of 1588 patients. *Rheumatology* 2017;56:1573–8.
- 14 Meeus M, Vervisch S, De Clerck LS, *et al.* Central sensitization in patients with rheumatoid arthritis: a systematic literature review. *Semin Arthritis Rheum* 2012;41:556–67.
- 15 Smolen JS, Breedveld FC, Schiff MH, *et al.* A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* 2003;42:244–57.
- 16 Aletaha D, Nell VPK, Stamm T, Uffmann M, *et al.* Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7:R796–806.
- 17 Boone NW, Sepriano A, van der Kuy P-H, *et al.* Routine assessment of patient index data 3 (RAPID3) alone is insufficient to monitor disease activity in rheumatoid arthritis in clinical practice. *RMD Open* 2019;5:e001050.
- 18 Boers M, Kirwan JR, Wells G, *et al.* Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol* 2014;67:745–53.
- 19 Bruce B, Fries JF. The health assessment questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14–18.
- 20 Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: identifying reversible and irreversible components. *Arthritis Rheum* 2006;54:2784–92.
- 21 Mease PJ, Etzel CJ, Huster WJ, *et al.* Understanding the association between skin involvement and joint activity in patients with psoriatic arthritis: experience from the Corrona registry. *RMD Open* 2019;5:e000867.
- 22 Landewé RBM. Overdiagnosis and overtreatment in rheumatology: a little caution is in order. *Ann Rheum Dis* 2018;77:1394–6.
- 23 Schoels MM, Smolen JS, Aletaha D. To DAPSA or not to DAPSA? that is not the question. *Ann Rheum Dis* 2019;78:e61.