Digital health technologies to strengthen patient-centred outcome assessment in clinical trials in inflammatory arthritis

Dr Dylan McGagh, BMBCh[1,2,3], Mr Kaiyang Song, BA[4], Dr Hang Yuan, DPhil[2,3], Dr Andrew P. Creagh, DPhil[5], Dr Sally Fenton, PhD[6,7], Prof Wan-Fai Ng, PhD[8,9,10], Ms Jennifer C. Goldsack, MBA[11], Prof William G Dixon, PhD[12,13,14], Prof Aiden Doherty, PhD[2,3], Prof Laura C Coates, PhD[1].

1. Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
2. Big Data Institute, University of Oxford, Oxford, UK
3. Nuffield Department of Population Health, University of Oxford, Oxford, UK
4. Oxford Medical School, Medical Sciences Division, University of Oxford, Oxford, UK
5. Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, UK
6. School of Sport, Exercise and Rehabilitation Science, University of Birmingham, Birmingham, UK
7. NIHR Birmingham Biomedical Research Centre, Birmingham, UK
8. HRB Clinical Research Facility, University College Cork, Cork, Ireland.
9. Translational and Clinical Research Institute, Newcastle University Faculty of Medical Sciences, Newcastle upon Tyne, UK
10. National Institute for Health and Care Research (NIHR), Newcastle Biomedical Research Centre & NIHR Newcastle Clinical Research Facility, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle Upon Tyne, UK.
11. The Digital Medicine Society (DiMe), Boston, MA, USA.
12. Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, The University of Manchester, Manchester, UK
13. NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

14. Department of Rheumatology, Salford Royal Hospital, Northern Care Alliance, Salford, UK

Corresponding author:

Dylan McGagh

dylan.mcgagh@ndorms.ox.ac.uk

Botnar Research Centre,
Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences,
Old Road,
Headington,
Oxford,
OX3 7LD,
United Kingdom

Summary

Common to all of the inflammatory arthritides, namely rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis, and juvenile idiopathic arthritis, is a potential for reduced mobility manifested through joint pain, swelling, stiffness and ultimately joint damage. Across these conditions, consensus has been reached on the need to capture outcomes related to mobility, such as functional capacity and physical activity as core domains in randomised controlled trials. Existing endpoints in use within these core domains rely wholly on self-reported questionnaires capturing patients' perception of their symptoms and activities which are subjective, inherently vulnerable to recall bias and fail to capture the granularity of fluctuations over time. There have been several early adopters integrating sensor-based digital health technology-derived endpoints to measure physical function and activity in randomised controlled trials for conditions including Parkinson's disease, Duchenne's muscular dystrophy, chronic obstructive pulmonary disease, and heart failure. Despite these applications, there have been no sensor-based digital health technology-derived endpoints in clinical trials recruiting patients with inflammatory arthritis. Borrowing from case studies across medicine, we outline the untapped opportunities and challenges in developing novel sensor-based digital health technology-derived endpoints which capture the symptoms and disease manifestations most relevant to patients with inflammatory arthritis.

3

Introduction

The inflammatory arthritides (IA) represent a collection of autoimmune rheumatic diseases including rheumatoid arthritis (RA), ankylosing spondyloarthritis (AS), psoriatic arthritis (PsA), and juvenile idiopathic arthritis (JIA).[1] These conditions are characterized by intra-articular inflammation which causes joint pain and stiffness as well as fatigue which can follow a fluctuating course between episodic clinical encounters. Over time, particularly if inflammation is not adequately controlled, patients with IA can progress to develop irreversible articular and peri-articular damage contributing to impaired joint mobility and limitations on individuals' ability to perform daily activities of living. This process can contribute to the marked overall functional and psychosocial disability experienced by people with IA.

The past 30 years have marked a substantial shift in the treatment and management of inflammatory arthritis (IA), profoundly improving the outlook for patients. With a growing number of treatments available that demonstrate the ability to improve disease activity and reduce inflammation, clinical trials are increasingly focused on the effect of different treatments on patient-relevant aspects such as mobility and physical function over time. This is also reflected in the emergence of the universal goal of treatment across IA to preserve physical function and social participation as a means of maximising health-related quality of life (HRQoL).[2,3] Despite this priority for both trials and clinical practice, assessment of such aspects of disease is currently reliant on self-reported questionnaires, which may lack sensitivity, are subject to recall bias and are burdensome for participants if used frequently. The over-reliance on such tools raises concern that some treatments may be incorrectly assessed as ineffective because of the deficiencies in, and poor performance of, current assessment instruments.

4

Across medicine, sensor-based digital health technologies (sDHTs) are emerging as a solution for the more objective assessment of patient-specific domains related to physical function and mobility, and its downstream relative, physical activity. sDHT-derived endpoints can capture a range of patient-centred concepts of interest and may broadly include data generated; a) while a patient performs a specific task designed to measure key symptoms, otherwise known as *active tasks* or b) through passive data acquisition monitoring human behaviours, such as physical activity or sleep, or physiological signs, such as heart rate or body temperature.[4] Active tasks can be administered using any sDHT, however, this is easiest on technologies with a screen for instructions, such as tablets and smartphones which have in-built gyroscopes and accelerometers measuring angular motion and velocity whilst passive data can be collected using a wide range of devices such as smartphones, smartwatches, wearables, and smart patches. In the context of patients with IA, active tasks could be deployed to capture individual joint range of motion tasks (i.e. a digitisation of the modified Schober's test in ankylosing spondylitis) or overall functional capacity through tests such as sit-to-stand or 6-Minute Walk Test. Passive wearable sensors, such as accelerometers, can directly measure movement behaviours, such as physical activity which is implicitly connected to musculoskeletal function. Sensor-based measurements can be collected remotely over a continuous period or at higher frequency and granularity, rather than restricting assessments only to intermittent clinic visits or through some remote electronic patient-reported outcome measures (ePROs).

We aim to present the current landscape of outcome assessment in IA trials before moving on to present the feasibility and validity of sensor-based digital health technologies in IA populations from existing observational research. We will examine several case studies of early adopters of sensor-based digital endpoints in trials across other areas of medicine to explore both the validity for inflammatory arthritis and regulatory steps towards digital

5

endpoint adoption in inflammatory arthritis trials. The final section will serve as a call to action and roadmap for the breadth of the rheumatology community to act in developing and integrating meaningful, effective and equitable sDHT-derived endpoints into RCTs.

Current deployment of patient-centred outcome measures in inflammatory arthritis trials

A significant contributor to the successful trial landscape in rheumatology has been the development of an extensive array of outcome measures. The Outcome Measures in Rheumatology (OMERACT) initiative has established international consensus on the core outcomes sets required in IA RCTs. Broadly, these core outcome sets reflect the need to capture either the underlying pathophysiological process (disease activity) or the manifestations of the condition on the individual (disease impact). In traditional efficacy trials, the gold standard objective disease activity outcome has been demonstrating differences in radiographic progression between placebo and treatment arms. There are ethical concerns about using radiographic progression with evidence to suggest irreversible joint damage and resulting functional limitation even from minor levels of radiographic progression.[5,6] In efforts to capture more of the heterogeneity of IA, OMERACT has endorsed several disease-specific composite disease activity measures.[7] There is emerging evidence suggesting rheumatology RCTs are at high risk of unblinding and there is often reported discordance between the physician and patient global domains, with patients reporting these measures do not accurately reflect their disease state.[8,9]

In efforts to increase patient-centricity in trial endpoints, OMERACT has unequivocally endorsed the use of measures that quantify the impact of treatment on physical function and health-related quality of life (HrQoL) across RCTs in IA. The only validated means of

6

measuring these core domains is through self-reported questionnaires, which are subject to several limitations, including notable floor and ceiling effects and a lack of sensitivity to change despite patients reporting disease improvement across various other domains.[10–12]

In patients with ankylosing spondylitis, there is only a moderate correlation between the OMERACT-endorsed physical function outcome, the Bath Ankylosing Spondylitis Functional Index (BASFI) and objective tests of physical function with patients appearing to incorporate exertion and pain levels in the self-assessment of physical function on the BASFI questionnaire.[13] The limitations of self-reported questionnaires of physical function have stimulated the development of more objective performance measures of physical function in AS.[14,15] The AS Performance Index (ASPI) is a clinic-based assessment which measures the time to perform three daily activities (bending to pick 6 pencils from the floor, putting on socks and standing up from the floor). The ASPI is still undergoing further validation prior to deployment in RCTs in AS; however, it has shown adequate to excellent intra-rater test-retest reliability, good responsiveness to treatment following TNF inhibition, and successfully measures different aspects of physical function when compared to the BASFI questionnaire.[14,15] These findings point to a benefit in measuring physical function through tests over simply collecting self-reported data. However, there are limitations with clinic-based performance measures with evidence showing that they capture maximal effort and are strongly influenced by participant motivation.[16]

What is the feasibility and validity of sensor-based digital endpoints in inflammatory arthritis?

Despite the lack of sensor-based digital health technology-derived endpoints in IA trials, there is a significant body of observational research investigating the use of sDHTs in a

7

variety of settings for patients with IA. These studies have investigated a range of sDHTs for measuring physical function and mobility, either deploying through active tasks completed by the participant (e.g. using a smartphone gyroscope which captures angular motion to measure specific joint range-of-motion or specific sit-to-stand or walking tasks to measure overall physical function), through passive wearable data (e.g. to measure physical activity level in a free-living environment), or a combination of these approaches and can serve as early milestones highlighting the potential validity of sDHT-derived endpoints in IA trials.

Active tasks

The PARADE pilot study utilised inbuilt gyroscopes in a smartphone to assess a range of active tasks related to a wrist range-of-motion task and a gait task to mobility at a multitude of affected joints in RA.[17] The authors found a relationship between reduced joint velocity in the range-of-motion task with higher pain intensity levels.[17] The ongoing Psorcast study, developed by Sage Bionetworks and the Psoriasis and Psoriatic Arthritis Clinics Multi-center Advancement Network (PPACMAN), is a prospective, observational cohort study recruiting participants with PsA to complete a range of active tasks on the Psorcast mobile application. This app leverages smartphone gyroscope and accelerometer-based active tasks in the form of the “Digital Jar Open” and 30-second walking task to monitor upper extremity function and overall physical function, respectively.[18] The “Digital Jar Open” provides a continuous variable outcome measure of overall upper limb angular rotation and as a single feature was able to distinguish participants with and without physician-assessed arm joint tenderness and/ or enthesitis and has the potential to remotely capture granular changes of functional impairment over time. There are other useful functional tests in inflammatory arthritis, such as the modified Schober’s test or hand grip strength, which could be digitised to form a remotely measurable active task, however, these are yet to be investigated. Active tasks can also include the use of electronic patient-reported outcomes (ePROs) which are often

8

digitisations of existing outcome measures. There is a large body of evidence supporting the feasibility, acceptability, and value of ePROs from observational studies across rheumatology, notably, highlighting the benefit of increasing the frequency of capturing symptoms over traditional measurements.[19] Most of this research is generated in the realm of remote monitoring vs standard clinical care and has been reviewed elsewhere in detail.[19]

Passively collected wearables data

Several other studies have investigated the relationship between accelerometer-assessed measures of physical activity with IA-specific factors including disease activity, physical function, and flare. Across PsA, two studies demonstrated an inverse relationship between disease activity and objectively measured physical activity volumes.[20,21] Hernandez-Hernandez found physical activity to be longitudinally responsive to fluctuations in disease activity across study visits.[20] In work by Prioreschi et al., treatment-naive RA patients who responded to DMARD therapy showed statistically significant improvements in physical activity volumes, with less sedentary and more light activity per day compared to non-responders.[22] In the ActConnect study, the authors leveraged physical activity data (steps per minute) to identify flares assessed with weekly self-reports collected from 20 healthy controls and 157 patients (83 with RA and 74 axial spondyloarthritis) over 3 months.[23] This study demonstrated the association between self-reported flare onset and modest reductions in physical activity volumes. The authors could predict a flare from accelerometer data alone with a sensitivity and specificity of 96% and 97%, respectively.[23] Collectively, these findings suggest that physical activity is closely related to the fluctuating course of disease activity in IA.

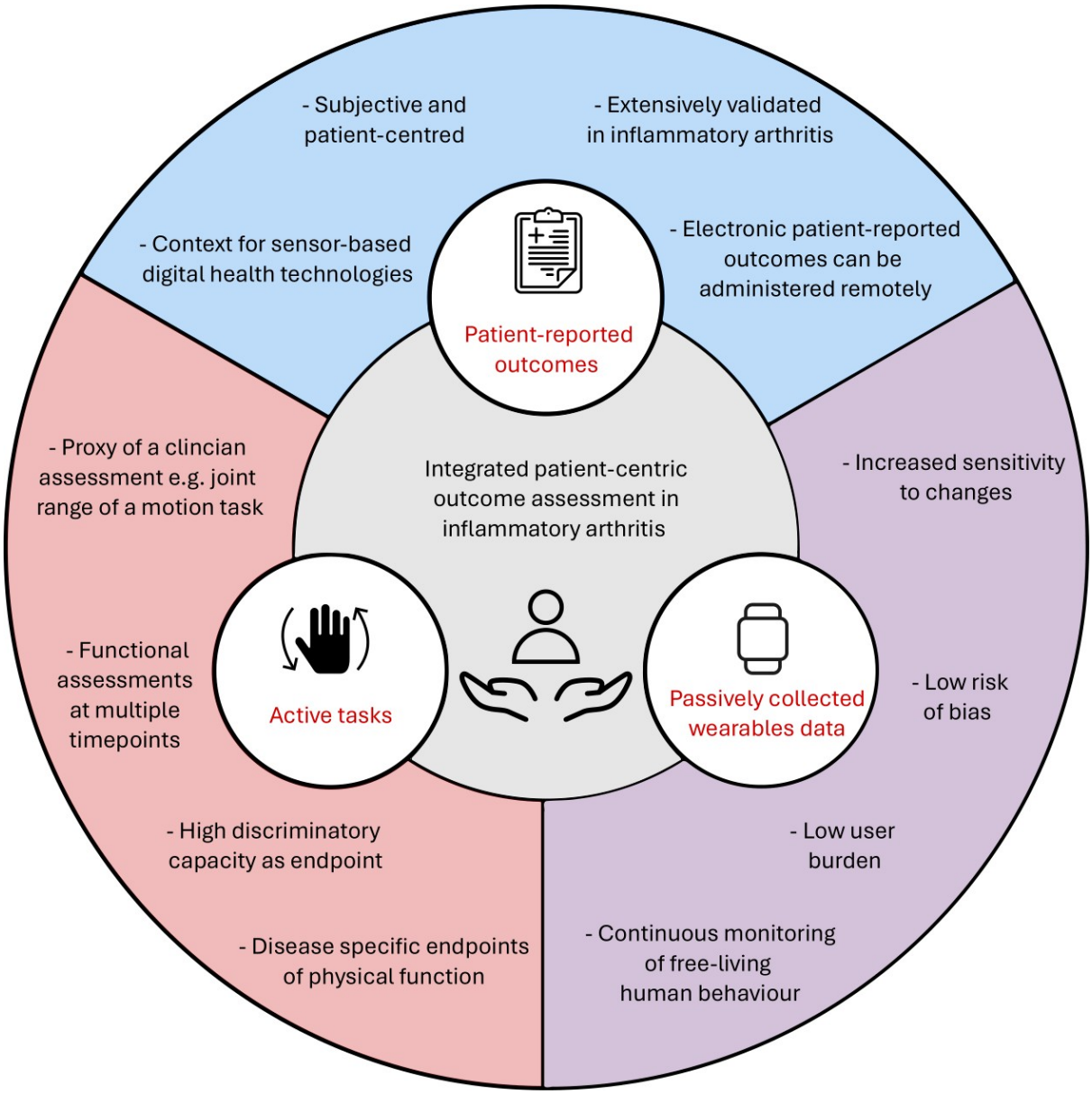
Integrated multi-modal approaches

9

The previously outlined approaches have tended towards testing a singular measure derived either from passively collected wearables data or active task data. An inherent limitation of any singular domain endpoint is the risk of confounding from factors outside the condition of interest. Overall physical activity may be affected both by the underlying pathophysiological process in inflammatory arthritis but also by acute intercurrent illness or comorbid diseases such as osteoarthritis, depression and anxiety and cardiometabolic diseases such as ischaemic heart disease or diabetes, all of which are associated with lower physical activity levels.[24] There is evidence emerging from the development of digital endpoints in COPD and Parkinson's disease around the utility of combining multiple data sources into a multi-modal composite digital endpoint to reduce the risk of confounding when compared to a singular measure.[25–28] To this end, the data garnered from an sDHT approach are placed synergistically alongside other patient-reported data to provide important context to the digital signal (as demonstrated in the context of IA in figure 1). There has been one data-intensive observational study testing the validity of a multi-modal DHT-based approach in capturing disease activity in rheumatology. The WeaRAble-PRO study investigated how standard PROs could be augmented with remote DHT sensor data, collected passively from smartwatches and via active tasks deployed on smartphones in patients with RA.[29] These tasks included ePROs as well as lying-to-standing mobility tasks using inbuilt gyroscopes to collect data on the longitudinal course of RA, accurately identifying participants in low and high disease activity states, as assessed by the RAPID-3 questionnaire.[29] A key insight from the WeaRAble-PRO study is the potential for sDHT-derived endpoints to complement and extend traditional PROs, offering a more nuanced view that encompasses both the subjective symptoms perceived by patients and the objective measurements captured by sDHT.[29] This perspective is crucial for patients with IA, where the impact extends beyond measurable

10

clinical signs to include fluctuating symptoms and higher comorbidity burden that can affect daily life and well-being



11

Figure 1 - Summary of key advantages of sensor-based digital health technology-derived endpoints and their integrated place alongside traditional patient-reported outcomes in the comprehensive assessment of patient-centred domains in inflammatory arthritis trials.

What is the uptake of digital endpoints outside of rheumatology?

There have been several early adopters of sensor-based digital health technology-derived endpoints in clinical trials across other areas of medicine. At the time of writing, the DiMe Library of Digital Endpoints reports 430 unique digital endpoints being used in industry-sponsored trials of new medical products. However, none have been deployed in any trial recruiting participants with a rheumatological condition.[30] Recent work by Masanneck et al. presented the growth trends of digital endpoints in use in neurology trials over the past decade.[31] The authors found that the relative frequency of clinical trials in neurology implementing digital health technology-derived endpoints has increased from 0.7% in 2010 to 11.4% in 2020 with the majority of these endpoints measuring motor function.[31] Across neurology, the poor performance of subjective questionnaire-based assessment tools for movement disorders has driven the development of digital endpoints for monitoring disease progression in Parkinson's disease and Duchenne's muscular dystrophy (DMD).[25,26,32] The first European Medicines Agency (EMA) approval of a digital endpoint for *stride velocity 95th centile* (SV95C) in patients with DMD was motivated by challenges presented by assessing physical function in clinic.[32] DMD is a progressive neuromuscular disease characterised by proximal muscle weakness and fatiguability with patients often experiencing significant burden to get to the clinic contributing to higher levels of fatigue confounding assessment hence the need for capturing the measure in a free-living environment. In rare

12

neurodegenerative conditions, such as Friedrich's Ataxia and amyotrophic lateral sclerosis, challenges around recruitment of sufficient sample sizes for RCTs has stimulated innovation in developing DHT-derived physical function endpoints which are more sensitive at detecting a minimal clinically important difference (MCID).[33,34] Akin to these neurological conditions, for patients with IA where mobility is so closely tied to functional capacity, these approaches have significant validity.

Beyond neurology, RCTs assessing interventions for heart failure have traditionally utilised efficacy endpoints that focus on mortality or hospitalisation. The availability of effective treatments has also driven an earlier shift towards head-to-head active comparator study designs when compared to rheumatology. As a result, this necessitates the design of extremely large and costly trials (median sample size of the 25 RCTs supporting the chronic heart failure guidelines was 2331 participants).[35] Furthermore, these endpoints give no information on how a patient's ability to function may change in response to an intervention. In response to this need to incorporate patient-centred perspectives on the symptoms and functional ability of patients in heart failure trials, The Heart Failure Collaboratory, a public-private partnership between regulators (US FDA), academia, payers, sponsors, and patients with HF was formed with a focus for integrating accelerometry-derived endpoints in HF trials.[36] A recent review by the Heart Failure Collaboratory highlights the 8 published RCTs using accelerometry for primary or secondary endpoints and a further six ongoing trials utilising this technology.[37] These trials have primarily focused on measuring change in daily step counts or global physical activity levels in response to treatment. The external validity of these studies to an IA population could be expanded due to the inherent relationship between joint pain and stiffness causing mobility limitations and subsequent reduced physical activity levels.

13

What are the regulatory considerations?

The development of digital endpoints is complex and includes several important steps leading to validation and approval. Progress towards implementation of sensor-based digital health technology-derived endpoints in clinical studies has been further reflected in recent regulatory trends, as seen in the first qualification of a wearable-derived digital endpoint by the EMA in the form of SV95C in patients with DMD,[32] whilst the FDA has finalized Guidance for industry regarding the use DHTs for remote data acquisition in clinical trials.[38] As these tools are being deployed to collect patient data to evaluate the safety and/or efficacy of a medicinal product in trials, regulatory bodies such as the EMA and FDA must ensure that measurements from these digital tools are clinically meaningful, representative, robust, and valid.[39] These regulatory bodies are both encouraging the formation of and actively participating in multistakeholder consortia with far-reaching domain expertise including patient perspective, clinicians, engineers, computer scientists, information governance, and industry to expedite DHT-derived endpoints that are valid and patient-centric.[40] Whilst some inconsistencies exist between the EMA and FDA regulation process, both bodies have endorsed the V3 Framework as a development process for any novel digital endpoint.[38,41,42] This first starts by framing key activities patients would wish to be able to complete but that their condition limits. In the context of patients with inflammatory arthritis, physical function or ability to perform daily ambulatory activities, is a key priority (Figure 2). From this, stakeholders can define a measurable outcome and identify a suitable DHT to capture this outcome. Once the need for a DHT is established, it is essential to demonstrate it is fit-for-purpose for gathering data in a clinical trial, for example showing that it is well tolerated by research and clinical volunteers.[42] Verification evaluates the performance of a sensor technology against pre-specified criteria to demonstrate the sample-level data generated is accurate e.g. raw data from accelerometer at specific measurement frequency, consistent over

14

time, and uniform across different environmental bench testing conditions. Analytical validation evaluates the performance of the sensor technology in human subjects, and its ability to measure, detect, or predict physiological or behavioural metrics against the pre-defined concept of interest. This step is agnostic of a specific clinical population of interest and can recruit healthy volunteers. The aim is to prove that the outcome of interest, for example, daily step count, is accurately captured when compared to a gold-standard measure, in this case, observer confirmed steps. Finally, clinical validation tests a digital measurement's predictive power in the context of its intended clinical use.[42] This step involves determining the optimal measurement window and frequency of sampling (e.g. days vs weeks) for a measure considering peak sensitivity, specificity, and reliability of a measure alongside user compliance and burden. This pipeline is an iterative, multi-study approach integrating patients, engineers, clinicians, industry partners, and regulatory bodies. This process can take several years to execute highlighting the need for timely engagement from the rheumatology community to realise the potential of these tools to capture meaningful measures for patients with IA in rheumatology RCTs.

15

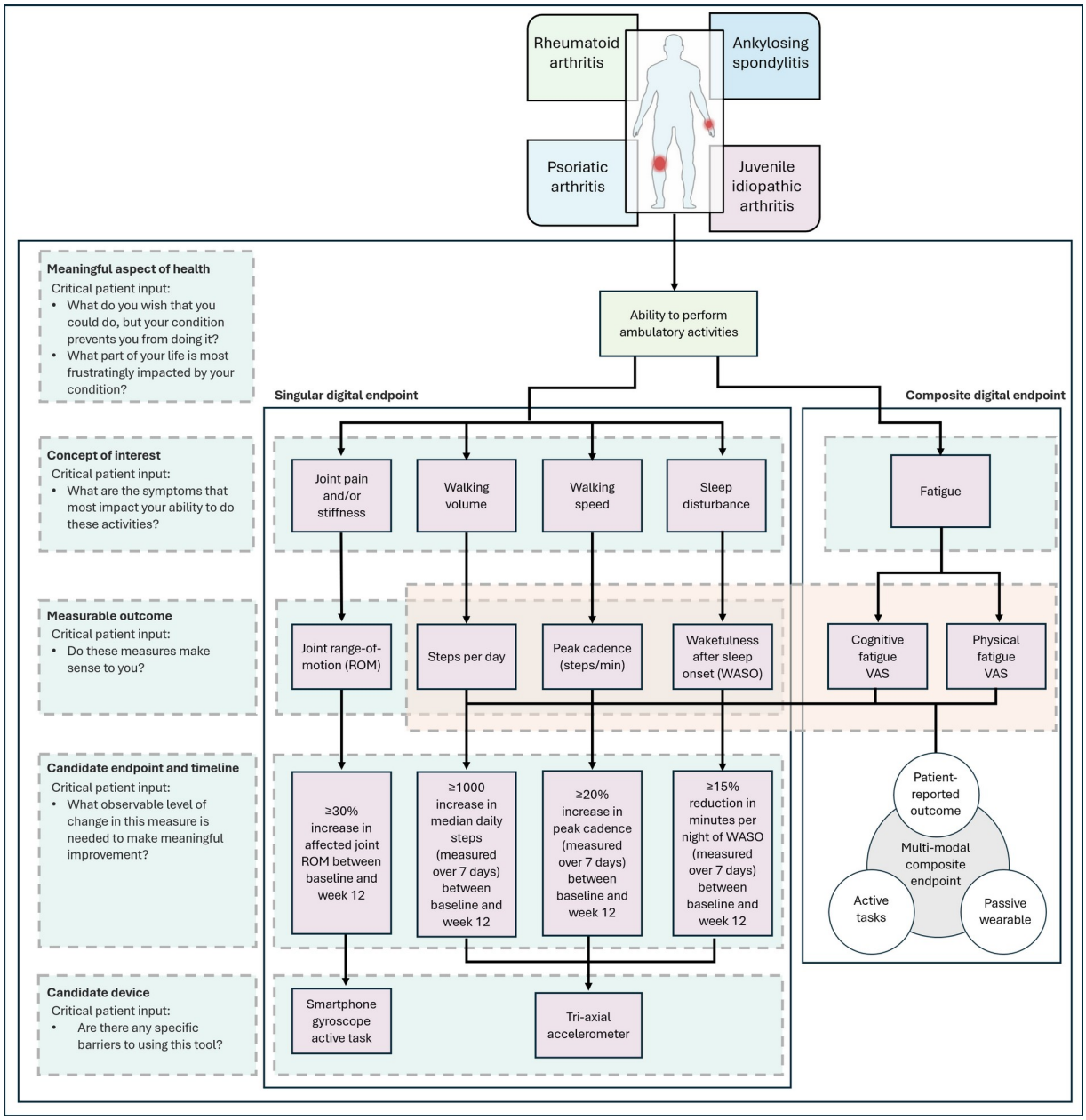


Figure 2 - A pipeline for the development patient-centric sensor-based digital endpoints for patients with inflammatory arthritis. (Adapted from Manta et al).[43] Measurable outcomes and resulting candidate endpoints outlined are hypothetical and purely illustrative of workflow. Specific validation studies are required to define the observable level of change as anchored to existing measures of physical function or health-related quality of life. Importantly, each step requires iterative patient input before progressing to the next step within a multi-study development pipeline. VAS; visual analogue scale.

Considerations regarding validity of digital endpoints in rheumatology

Whilst the implementation of sensor-based digital health technologies offers a means of obtaining objective, high-resolution data from patients captured during their daily lives, their

16

use also poses some challenges. Large and diverse validation datasets are needed to avoid algorithmic biases and in turn the inaccurate collection and interpretation of digital health data from certain demographic groups.[44] Furthermore, clinical validation studies for sDHT-endpoints in IA anchored to existing disease activity and physical function measures need to recruit groups of participants that are diverse and representative.Other important issues include variable data quality, often stemming from improper wearing and/or a lack of compliance with DHT hardware, as well as the psychological impact of wearable devices. Sleep data from sensory devices can skew patients''perception of their quality of sleep, and in turn, reinforce sleep-related anxiety (i.e. orthosomnia).[45] In the context of inflammatory arthritis where individuals may have disease manifestations affecting their fine-motor movements of the hands, it is imperative to ensure that any adjustments to the sensor or display are reasonably developed with these considerations so as to reduce the risk of digital exclusion. In pursuit of developing robust and valid digital endpoints in IA, special attention should be paid to whether candidate measurements are specific IA alone rather than being global indicators of illness. In developing sDHT-endpoints in juvenile idiopathic arthritis, understanding the important confounding influence of age on physical activity level needs to be considered and any endpoint needs to be validated within specific age subgroups to account for changes in stride length and velocity. Altogether, the concerns outlined exemplify the importance of well-designed technical and clinical validation studies for a given DHT, involving patients with diverse demographic characteristics. Trials in inflammatory arthritis have systemically underrepresented and underrecruited patients from racial and ethnic minority groups.[46] Whilst not explored in detail in this review, these endpoints offer opportunities for remote data capture and a shift towards decentralised trial designs. Following a patient-centric approach to endpoint development and integrating patient partners at each stage of the development pipeline (figure 2) offers the opportunity to reduce

17

biases and increase equity in trials for people with inflammatory arthritis. This process can also be formalised through specific Health Equity Impact Assessment (HEIA) processes which aim to identify and address any risk of digital exclusion.[47,48]

What role does machine learning play?

The field of digital endpoint adoption in clinical trials is rapidly evolving. Recent regulatory guidance from the FDA on the use of DHTs in endpoint development did not address the role of machine learning. Despite this, machine learning is becoming ever more widely used to interpret the highly granular data collected using sensor-based digital health technologies. Through machine learning, it becomes possible to discern intricate patterns and relationships within the data, which might remain undetected with conventional statistical approaches. By learning from large volumes of data, foundation models have been shown to improve the reliability of digital measures across populations, devices, and modalities, critical for the deployment of DHTs in clinical trials.[49] Particularly relevant is the value of machine learning in multi-modal sDHT approaches that enable the development of novel digital markers,[33,34] and more precise monitoring of disease progression.[29] Machine learning techniques allow for the integration and analysis of diverse data sources, including active tasks, accelerometer-measured physical activity, and patient-reported outcomes. This could be particularly important in the context of developing composite sDHT-derived endpoints, akin to a digital endpoint of fatigue proposed in figure 2. Whilst machine learning holds immense promise in this field, there are important considerations to ensure its effective implementation. Key considerations include promoting model explainability, ensuring transparency and reproducibility, establishing robust governance infrastructures, training models using relevant and representative clinical validation datasets and addressing bias.[50,51]

18

What is the overall set of actions we need to enact as a community to realise this vision in an IA population?

Through OMERACT there exists an infrastructure that integrates academics, industry, regulators, and patient partners to develop and endorse core outcomes for implementation in rheumatology RCTs. A specific OMERACT Digital Working Group with the view to considering meaningful aspects of health across IA that could be measured using sensor-based digital health technologies is imperative to integrate these outcomes in IA trials. There have been strides taken to create core outcome sets related to digital measures of physical activity and mobility which would serve as a focus for such a group.[52] The development of sDHT-derived endpoints is a lengthy and intricate process, underscoring the urgency for immediate action to initiate this vital journey toward revolutionizing patient-centric outcomes in clinical trials in inflammatory arthritis.

In the pursuit of developing digital endpoints in rheumatology, the model set out by the ambitious IDEA-FAST and MOBILISE-D consortia should be followed closely.[53,54] These consortia consist of patient partners, academia, and industry stakeholders aiming to develop cross-disease endpoints based on meaningful aspects of health which are conserved across a multitude of diseases, for example, fatigue being relevant in neurodegenerative diseases such as Parkinson's disease as well as immune-mediated inflammatory diseases such as RA or Crohn's disease.[53] This cross-disease approach aims to accelerate device verification and endpoint validation steps with both IDEA-FAST and MOBILISE-D having received positive support for this approach from the EMA.[53,55] Further to this cross-disease approach, recent work by Bertha et al. proposes a framework building on the V3 verification and validation pipeline which permits leveraging verification evidence from a prior related disease area and

19

application to a new related use case.[56] The work from IDEA-FAST, MOBILISE-D, and Bertha et al. have significant relevance for the rheumatology community given there is a high degree of harmony in patient-centred concepts of interest across each of the IA, potentially enabling cross-disease validation studies of DHT-derived endpoints which could accelerate the development process.

Conclusion

The integration of sensor-based digital health technology-derived endpoints holds immense promise for rheumatology, offering the potential to revolutionize assessment in IA trials. Whilst the journey towards digital endpoint development in rheumatology may be challenging, it is paved with a wealth of valuable case studies from across medicine. Drawing from these exemplars, there is a robust foundation to build upon, offering a clear pathway for the successful integration of digital endpoints supporting shorter, more cost-effective, and patient-centred RCTs that result in new therapies that address the symptoms most important to all individuals with IA.

Search strategy and selection criteria

This narrative review was supported by a literature search in PubMed and Google Scholar, focusing on publications in English from 2010 to 2024. Search terms included “inflammatory arthritis”, “rheumatoid arthritis”, “ankylosing spondylitis”, “psoriatic arthritis”, “juvenile

20

idiopathic arthritis", "outcome measures", "patient-reported outcome measures", "OMERACT", "digital health technology", "digital endpoint", "digital health" and "randomized controlled trials". Additional publications were identified from references using a pearl-growing approach.

Conflicts of interest statement

HY has received a personal payment via University of Oxford for software licenses he has developed. APC; this work was conducted while APC was an employee of University of Oxford. During this time, APC was partially funded by GSK plc. and has worked on projects related to GSK. APC has never had a contract of employment, owned any stocks or shares, equity, or had a named position on a company board associated with GSK. GSK has no involvement or affiliation with this review work, in any kind. On submission of this work, APC is currently an employee and shareholder of Sanome Ltd. Sanome has no involvement or affiliation with this review work, in any kind. APC was involved in initial idea generation, conceptual work of this review, while he was still employed by the University of Oxford, which predates his employment with Sanome. The author declares that there is no direct or indirect conflict of interest between their previous affiliation with the University of Oxford and their current employment at Sanome regarding the material discussed in this manuscript. JG sits as an unpaid volunteer member on the Board of Directors for the Digital Medicine Society and Coalition for Health AI. WGD has received consultancy fees from Google unrelated to this work. AD is supported by the Wellcome Trust [223100/Z/21/Z], Novo Nordisk, Swiss Re, the British Heart Foundation Centre of Research Excellence (grant number RE/18/3/34214), and Health Data Research UK, an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities. AD's research group has also received funding from GlaxoSmithKline (GSK). AD has received personal payments via

21

University of Oxford for a) consulting services provided to Harvard University and University of Wisconsin. AD has received a personal payment via University of Oxford for software licenses he has developed. LCC has received research grants from AbbVie, Amgen, Celgene, Eli Lilly, Janssen, Novartis, Pfizer and UCB. LCC has received consulting fees from AbbVie, Amgen, Bristol Myers Squibb, Celgene, Eli Lilly, Gilead, Galapagos, Janssen, Moonlake, Novartis, Pfizer and UCB. LCC has received payment for lectures, presentations AbbVie, Amgen, Biogen, Celgene, Eli Lilly, Galapagos, Gilead, GSK, Janssen, Medac, Novartis, Pfizer and UCB. The other authors declared no conflicts of interest.

Author contributions

DMcG - Conceptualisation, Visualization, Writing – original draft, Writing – review & editing

KS - Conceptualisation, Writing – original draft, Writing – review & editing

HY - Conceptualisation, Visualization, Writing – original draft, Writing – review & editing

22

APC - Conceptualisation, Writing – original draft, Writing – review & editing

SF- Conceptualisation, Visualization, Writing – review & editing

WFN - Conceptualisation, Writing – review & editing

JCG - Conceptualisation Writing – review & editing

WGD - Conceptualisation, Writing – review & editing

AD - Conceptualisation, Writing – review & editing, Supervision

LCC – Conceptualisation, Writing – review & editing, Supervision

References

1 Ledingham J, Snowden N, Ide Z. Diagnosis and early management of inflammatory arthritis. *BMJ (Online)* 2017; **358**. DOI:10.1136/bmj.j3248.

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**. DOI:10.1136/ard.2009.123919.

3 Smolen JS, Schöls M, Braun J, *et al.* Treating axial spondyloarthritis and peripheral spondyloarthritis, especially psoriatic arthritis, to target: 2017 update of recommendations by an international task force. *Ann Rheum Dis* 2018; **77**: 3–17.

4 Taylor KI, Staunton H, Lipsmeier F, Nobbs D, Lindemann M. Outcome measures based on digital health technology sensor data: data- and patient-centric approaches. *NPJ Digit Med* 2020; **3**: 97.

5 van der Heijde D, Landewé R, van Vollenhoven R, Fatenejad S, Klareskog L. Level of radiographic damage and radiographic progression are determinants of physical function: a longitudinal analysis of the TEMPO trial. *Ann Rheum Dis* 2008; **67**: 1267–70.

23

van der Heijde D, Landewé R. Should radiographic progression still be used as outcome in RA? *Clinical Immunology* 2018; **186**: 79–81.

Day J, Antony A, Tillett W, Coates LC. The state of the art—psoriatic arthritis outcome assessment in clinical trials and daily practice. *Lancet Rheumatol* 2022; **4**: e220–8.

Bruggemeyer C, Nepal D, Putman M. Unintentional unblinding in rheumatic disease trials. *Lancet Rheumatol* 2023; **5**: e633–6.

Desthieux C, Hermet A, Granger B, Fautrel B, Gossec L. Patient-Physician Discordance in Global Assessment in Rheumatoid Arthritis: A Systematic Literature Review With Meta-Analysis. *Arthritis Care Res (Hoboken)* 2016; **68**: 1767–73.

Maska L, Anderson J, Michaud K. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). *Arthritis Care Res (Hoboken)* 2011; **63 Suppl 11**: S4-13.

Orbai A-M, Ogdie A. Patient-Reported Outcomes in Psoriatic Arthritis. *Rheum Dis Clin North Am* 2016; **42**: 265–83.

Højgaard P, Klokker L, Orbai A-M, *et al.* A systematic review of measurement properties of patient reported outcome measures in psoriatic arthritis: A GRAPPA-OMERACT initiative. *Semin Arthritis Rheum* 2018; **47**: 654–65.

van Weely SFE, van Denderen JC, Steultjens MPM, *et al.* Moving instead of asking? Performance-based tests and BASFI-questionnaire measure different aspects of physical function in ankylosing spondylitis. *Arthritis Res Ther* 2012; **14**. DOI:10.1186/ar3765.

24

54 van Weely SFE, Dekker J, Steultjens MPM, *et al.* Objective evaluation of physical
55 functioning after tumor necrosis factor inhibitory therapy in patients with ankylosing
56 spondylitis: a selection of 3 feasible performance-based tests. *J Rheumatol* 2015; **42**: 623–9.
57 van Bentum RE, Ibáñez Vodnizza SE, Poblete de la Fuente MP, *et al.* The Ankylosing
58 Spondylitis Performance Index: Reliability and Feasibility of an Objective Test for Physical
59 Functioning. *J Rheumatol* 2020; **47**: 1475–82.
60 Alfano LN, Lowes LP, Berry KM, *et al.* T.P.1: Pilot study evaluating motivation on the
61 performance of timed walking in boys with Duchenne muscular dystrophy. *Neuromuscular*
62 *Disorders* 2014; **24**: 860.
63 Hamy V, Garcia-Gancedo L, Pollard A, *et al.* Developing Smartphone-Based Objective
64 Assessments of Physical Function in Rheumatoid Arthritis Patients: The PARADE Study.
65 *Digit Biomark* 2020; **4**: 26–44.
66 Webster DE, Haberman RH, Perez Chada LM, *et al.* Clinical validation of digital biomarkers
67 and machine learning models for remote measurement of psoriasis and psoriatic arthritis.
68 *medRxiv* 2022.
69 Marques A, Bosch P, de Thurah A, *et al.* Effectiveness of remote care interventions: a
70 systematic review informing the 2022 EULAR Points to Consider for remote care in
71 rheumatic and musculoskeletal diseases. *RMD Open* 2022; **8**. DOI:10.1136/rmdopen-2022-
72 002290.
73 Hernández-Hernández MV, Sánchez-Pérez H, Luna-Gómez C, Ferraz-Amaro I, Díaz-
74 González F. Impact of Disease Activity on Physical Activity in Patients With Psoriatic
75 Arthritis. *Arthritis Care Res (Hoboken)* 2021; **73**. DOI:10.1002/acr.24422.
76 McGagh D, McGowan N, Hinds C, Saunders KEA, Coates LC. Actigraphy-derived physical
77 activity levels and circadian rhythm parameters in patients with psoriatic arthritis:

25

relationship with disease activity, mood, age and BMI. *Ther Adv Musculoskelet Dis* 2023; **15**: 1759720X231174989.

Prioreschi A, Hodkinson B, Tikly M, Mcveigh JA. Changes in physical activity measured by accelerometry following initiation of dmard therapy in rheumatoid arthritis. *Rheumatology (United Kingdom)* 2014; **53**. DOI:10.1093/rheumatology/ket457.

Gossec L, Guyard F, Leroy D, *et al.* Detection of Flares by Decrease in Physical Activity, Collected Using Wearable Activity Trackers in Rheumatoid Arthritis or Axial Spondyloarthritis: An Application of Machine Learning Analyses in Rheumatology. *Arthritis Care Res (Hoboken)* 2019; **71**. DOI:10.1002/acr.23768.

Barker J, Smith Byrne K, Doherty A, *et al.* Physical activity of UK adults with chronic disease: cross-sectional analysis of accelerometer-measured physical activity in 96 706 UK Biobank participants. *Int J Epidemiol* 2019; **48**: 1167–74.

Jha A, Menozzi E, Oyekan R, *et al.* The CloudUPDRS smartphone software in Parkinson's study: cross-validation against blinded human raters. *NPJ Parkinsons Dis* 2020; **6**: 36.

Lipsmeier F, Taylor KI, Postuma RB, *et al.* Reliability and validity of the Roche PD Mobile Application for remote monitoring of early Parkinson's disease. *Sci Rep* 2022; **12**: 12081.

Jha A, Espay AJ, Lees AJ. Digital Biomarkers in Parkinson's Disease: Missing the Forest for the Trees? *Mov Disord Clin Pract* 2023; **10**: S68–72.

Garcia-Aymerich J, Puhan MA, Corriol-Rohou S, *et al.* Validity and responsiveness of the Daily- and Clinical visit-PROactive Physical Activity in COPD (D-PPAC and C-PPAC) instruments. *Thorax* 2021; **76**: 228–38.

Creagh AP, Hamy V, Yuan H, *et al.* Digital health technologies and machine learning augment patient reported outcomes to remotely characterise rheumatoid arthritis. *NPJ Digit Med* 2024; **7**: 33.

26

Digital Medicine Society (DiMe). Digital Medicine Society (DiMe) Library of Digital Endpoints. https://dimesociety.org/get-involved/library-of-digital-endpoints/. 2023; published online Aug.

Masanneck L, Gieseler P, Gordon WJ, Meuth SG, Stern AD. Evidence from ClinicalTrials.gov on the growth of Digital Health Technologies in neurology trials. *NPJ Digit Med* 2023; **6**. DOI:10.1038/s41746-023-00767-1.

Servais L, Camino E, Clement A, *et al.* First Regulatory Qualification of a Novel Digital Endpoint in Duchenne Muscular Dystrophy: A Multi-Stakeholder Perspective on the Impact for Patients and for Drug Development in Neuromuscular Diseases. *Digit Biomark* 2021; **5**: 183–90.

Gupta AS, Patel S, Premasiri A, Vieira F. At-home wearables and machine learning sensitively capture disease progression in amyotrophic lateral sclerosis. *Nat Commun* 2023; **14**: 5080.

Ricotti V, Kadirvelu B, Selby V, *et al.* Wearable full-body motion tracking of activities of daily living predicts disease trajectory in Duchenne muscular dystrophy. *Nat Med* 2023; **29**. DOI:10.1038/s41591-022-02045-1.

Docherty KF, Campbell RT, Jhund PS, Petrie MC, McMurray JJ V. How robust are clinical trials in heart failure? *Eur Heart J* 2017; **38**: 338–45.

Psotka MA, Abraham WT, Fiuzat M, *et al.* Functional and Symptomatic Clinical Trial Endpoints: The HFC-ARC Scientific Expert Panel. *JACC Heart Fail* 2022; **10**: 889–901.

Anchouche K, Elharram M, Oulousian E, *et al.* Use of Actigraphy (Wearable Digital Sensors to Monitor Activity) in Heart Failure Randomized Clinical Trials: A Scoping Review. *Can J Cardiol* 2021; **37**: 1438–49.

27

U.S. Food and Drug Administration (FDA). Digital Health Technologies for Remote Data Acquisition in Clinical Investigations. 2023.

Colloud S, Metcalfe T, Askin S, *et al.* Evolving regulatory perspectives on digital health technologies for medicinal product development. *NPJ Digit Med* 2023; **6**: 56.

Stephenson D, Alexander R, Aggarwal V, *et al.* Precompetitive Consensus Building to Facilitate the Use of Digital Health Technologies to Support Parkinson Disease Drug Development through Regulatory Science. *Digit Biomark* 2020; **4**: 28–49.

Clinical Trials Transformation Initiative (CTTI). Developing Novel Endpoints Generated by Digital Health Technology for Use in Clinical Trials. 2022.

Goldsack JC, Coravos A, Bakker JP, *et al.* Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *NPJ Digit Med* 2020; **3**: 55.

Manta C, Patrick-Lake B, Goldsack JC. Digital Measures That Matter to Patients: A Framework to Guide the Selection and Development of Digital Measures of Health. Digit Biomark. 2020; **4**. DOI:10.1159/000509725.

Zinzuwadia A, Singh JP. Wearable devices-addressing bias and inequity. *Lancet Digit Health* 2022; **4**: e856–7.

Baron KG, Abbott S, Jao N, Manalo N, Mullen R. Orthosomnia: Are Some Patients Taking the Quantified Self Too Far? *J Clin Sleep Med* 2017; **13**: 351–4.

Strait A, Castillo F, Choden S, *et al.* Demographic Characteristics of Participants in Rheumatoid Arthritis Randomized Clinical Trials: A Systematic Review. *JAMA Netw Open* 2019; **2**: e1914745.

Were MC, Sinha C, Catalani C. A systematic approach to equity assessment for digital health interventions: case example of mobile personal health records. *J Am Med Inform Assoc* 2019; **26**: 884–90.

28

Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. *NPJ Digit Med* 2022; **5**: 119.

Yuan H, Chan S, Creagh AP, *et al.* Self-supervised learning for human activity recognition using 700,000 person-days of wearable data. *NPJ Digit Med* 2024; **7**: 91.

U.S. Food and Drug Administration. Good Machine Learning Practice for Medical Device Development: Guiding Principles. 2021.

European Medicines Agency (EMA). Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle. 2023.

Digital Medicine Society (DiMe). The core set of digital measures of physical activity. https://datacc.dimesociety.org/digital-measures-physical-activity/all-core-measures/. 2024; published online Feb 15.

Nobbs D, Piwko W, Bull C, *et al.* Regulatory Qualification of a Cross-Disease Digital Measure: Benefits and Challenges from the Perspective of IMI Consortium IDEA-FAST. *Digit Biomark* 2023; **7**: 132–8.

Mikolaizak AS, Rochester L, Maetzler W, *et al.* Connecting real-world digital mobility assessment to clinical outcomes for regulatory and clinical endorsement-the Mobilise-D study protocol. *PLoS One* 2022; **17**: e0269615.

European Medicines Agency (EMA). Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. 2020.

Bertha A, Alaj R, Bousnina I, *et al.* Incorporating digitally derived endpoints within clinical development programs by leveraging prior work. *NPJ Digit Med* 2023; **6**: 139.

29