# Evaluation of a Musculoskeletal Digital Assessment Routing Tool [DART]: A Crossover Non-Inferiority Randomized Pilot Trial

Cabella Lowe, Ruth Sephton, William Marsh, Dylan Morrissey

Submitted to: JMIR Formative Research
on: February 02, 2024

# *Table of Contents*

# Evaluation of a Musculoskeletal Digital Assessment Routing Tool [DART]: A Crossover Non-Inferiority Randomized Pilot Trial

Cabella Lowe[1] MSc; Ruth Sephton[2] MSc; William Marsh[3] PhD; Dylan Morrissey[1, 4] PhD

[1]Centre for Sports & Exercise Medicine William Harvey Research Institute Queen Mary University of London London GB
[2]St Helens Musculoskeletal Physiotherapy Service Mersey Care NHS Foundation Services St Helens GB
[3]Machine Intelligence and Decision Support [MInDS] Research Group School of Electronic Engineering and Computer Science and Digital Environment Research Institute Queen Mary University of London London GB
[4]Department of Physiotherapy Barts Health NHS Trust London GB

**Corresponding Author:**
Cabella Lowe MSc
Centre for Sports & Exercise Medicine
William Harvey Research Institute
Queen Mary University of London
Mile End Hospital
Bancroft Road
London
GB



## *Abstract*

**Background:** Musculoskeletal conditions account for 16% of global disability, resulting in a negative effect on millions of patients and increasing demand for healthcare utilization. Introduction of digital technologies to improve health care outcomes and efficiency have been prioritized . We have developed a musculoskeletal Digital Assessment Routing Tool (DART), enabling patients to self-assess and be directed to the right care which requires validation prior to implementation. Such innovations are rarely rigorously tested in clinical trials - considered the gold standard for evaluating safety and efficacy. This pilot study is a precursor to a trial assessing DART performance with a physiotherapist-led triage assessment.

**Objective:** To evaluate trial design, assess procedures and collect exploratory data to assess feasibility of delivering an adequately powered, definitive randomized trial, assessing DART safety and efficacy in an NHS primary care setting.

**Methods:** This 8-week crossover, non-inferiority pilot trial utilizing an Integrated Knowledge Translation approach, took place in an NHS England primary care practice. Participants were patients over 18 years, registered with the practice and seeking assessment for a musculoskeletal condition. All participants completed a DART assessment and the history-taking element of a face-to-face physio-led triage in a randomized order. The primary outcome was the agreement between DART and physiotherapist triage recommendation to condition management pathway. Data were collected allowing analysis of participant recruitment and retention, randomization, blinding, study burden and potential barriers to intervention delivery. Participant satisfaction with using DART was measured using the System Usability Scale.

**Results:** 129 patients were invited to participate, with 60% (78/129) meeting the inclusion criteria and being randomized into each intervention arm (39/39). There were no dropouts and data were analyzed for all 78 participants. Agreement between physiotherapist and DART across all participants and all primary triage outcome was 32/78, 41% (95% CI 22 to 45), ICC=0.37 (95% CI = 0.16-0.55), indicating that the reliability of DART was poor to moderate. Feedback from the clinical service team led to adjusted analysis yielding 61/78, 78% (CI 47 to 78), ICC=0.57 with a 95% confidence interval = 0.40-0.70. Participant satisfaction was measured quantitively using amalgamated System Usability Scale scores (n=78, mean score = 84.0, 90% CI = ±2.94), equating to a “Excellent” system. There were no study incidents, and trial burden was acceptable.

**Conclusions:** This pilot highlighted the well documented complexity of assessing safety and effectiveness of a digital triage system. In response to this, amendments to the study protocol are proposed to improve validity of the main trial. Completion of a consensus study is recommended to inform what constitutes an acceptable noninferiority margin and subsequent calculation of the full trial sample size. It is concluded an adequately powered definitive noninferiority randomized controlled trial is feasible. Clinical Trial: Clinicaltrials.gov NCT04904029, http://clinicaltrials.gov/ct2/show/NCT04904029

(JMIR Preprints 02/02/2024:56715)
DOI: https://doi.org/10.2196/preprints.56715

## Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

# Original Manuscript

# Evaluation of a Musculoskeletal Digital Assessment Routing Tool [DART]: A Crossover Non-Inferiority Randomized Pilot Trial

Cabella Lowe[1], Ruth Sephton[2], William Marsh[3], Dylan Morrissey[1]
**[1] Centre for Sports & Exercise Medicine, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom**
**[2] St Helens Musculoskeletal Physiotherapy Service, Mersey Care NHS Foundation Services, St Helens, United Kingdom**

**[3]Machine Intelligence and Decision Support [MInDS] Research Group, School of Electronic Engineering and Computer Science and Digital Environment Research Institute, Queen Mary University of London, London, United Kingdom**

Corresponding Author:
Cabella Lowe, MSc, MCSP, MMACP
Centre for Sports & Exercise Medicine
Queen Mary University of London
London
United Kingdom
Phone: 44 7976 315105
Email: c.lowe@qmul.ac.uk



## Abstract

**Background:** Musculoskeletal conditions account for 16% of global disability, resulting in a negative effect on patients and increasing demand for healthcare utilization. Triage directing patients to appropriate level intervention improving health outcomes and efficiency have been prioritized . We developed a musculoskeletal Digital Assessment Routing Tool (DART) mHealth system, which requires evaluation prior to implementation. Such innovations are rarely rigorously tested in clinical trials - considered the gold standard for evaluating safety and efficacy. This pilot study is a precursor to a trial assessing DART performance with a physiotherapist-led triage assessment.

**Objective:** To evaluate trial design, assess procedures and collect exploratory data to establish feasibility of delivering an adequately powered, definitive randomized trial, assessing DART safety and efficacy in an NHS primary care setting.

**Methods:** A crossover, non-inferiority pilot trial utilizing an Integrated Knowledge Translation approach within an NHS England primary care setting. Participants were patients seeking assessment for a musculoskeletal condition, completing a DART assessment and the history-taking element of a face-to-face physio-led triage in a randomized order. The primary outcome was agreement between DART and physiotherapist triage recommendation. Data allowed analysis of participant recruitment and retention, randomization, blinding, study burden and potential barriers to intervention delivery. Participant satisfaction was measured using the System Usability Scale.

**Results:** Over 8 weeks, 129 patients were invited to participate. Of these, 92% (119/129) proceeded to eligibility assessment, with 60% (78/129) meeting the inclusion criteria and being randomized into each intervention arm (39/39). There were no dropouts and data were analyzed for all 78 participants. Agreement between physiotherapist and DART across all participants and all primary triage outcome was 32/78, 41% (95% CI 22 to 45), ICC=0.37 (95%

CI = 0.16-0.55), indicating that the reliability of DART was poor to moderate. Feedback from the clinical service team led to adjusted analysis yielding 61/78, 78% (CI 47 to 78), ICC=0.57 with a 95% confidence interval = 0.40-0.70. Participant satisfaction was measured quantitively using amalgamated System Usability Scale scores (n=78, mean score = 84.0, 90% CI = ±2.94), equating to a "Excellent" system. There were no study incidents, and trial burden was acceptable.

**Conclusions:** Physiotherapist-DART agreement of 78%, with no adverse triage decisions and high patient satisfaction, was sufficient to conclude DART had potential to improve the MSK pathway. Study validity was enhanced by recruitment of real-world patients and using an Integrated Knowledge Translation approach. Completion of a context-specific consensus process is recommended to provide definitive definitions of safety criteria, range of appropriateness, non-inferiority margin and sample size. This pilot demonstrated an adequately powered definitive trial is feasible, which would provide evidence of DART safety and efficacy, ultimately informing potential for DART implementation.
Registration Clinicaltrials.gov NCT04904029,

http://clinicaltrials.gov/ct2/show/NCT04904029

The protocol for this trial can be accessed from:
**Lowe C**, Hanuman Sing H, Marsh W, Morrissey D, Validation of a Musculoskeletal Digital Assessment Routing Tool: Protocol for a Pilot Randomized Crossover Noninferiority Trial JMIR Res Protoc **2021;10[12]:e31541 URL: https://www.researchprotocols.org/2021/12/e31541 DOI: 10.2196/31541**
Keywords: mHealth; mobile health; eHealth; digital health; digital technology; musculoskeletal;

triage; physiotherapy triage

## Introduction

### Background

Musculoskeletal (MSK) conditions are a global epidemic, prevalent across all ages and increasing rapidly [1–3], being associated with increased life expectancy and reduced activity [4,5]. In the United Kingdom(UK) MSK conditions pose a financial and societal challenge, costing over £4.76 billion of the UK National Health Service (NHS) resources and utilising up to one in three primary care physician visits annually [6,7]. Patients utilize more healthcare and generate higher costs if they must wait longer for assessment and treatment [8, 9] with longer waiting times potentially leading to detrimental effects on pain, disability, and quality of life for waiting patients [8,9], as well as increasing their risk of chronic health disease [10]. "Getting It Right First Time" (GIRFT) by directing patients to the correct level of intervention at first point of contact, is considered key in improving condition outcomes and reducing unwarranted variation in clinical pathways, such as unnecessary secondary care consultations and investigations [11].

Remote physiotherapist led MSK triage services are widely utilized within the UK NHS and private sectors and have potential to reduce waiting times, MSK caseload and cost across the pathway [12,13, 14]. However, the principal rate-limiting factor on the ability of services to increase activity and treat more patients is the availability of staff [15].
It has been suggested Mobile Health (mHealth) technology could provide a cost-effective

alternative to physiotherapist-led remote triage for improving healthcare delivery [16, 17], with recent advances being made in digital primary care triage applications [18,19]. Using a digital triage tool has the potential to screen for conditions requiring emergency or urgent care, whilst directing less complex or urgent presentations to routine physiotherapy or supported self-management, thereby maximizing utilization of highly skilled clinicians' time.

However, a web-based triage platform directing patients with MSK presentations to an appropriate level of care requires robust testing and validation prior to implementation [20-28]. To date there is limited evidence regarding the use of web-based or digital triage platforms for MSK conditions specifically, with most investigations focused on the performance of generic symptom checkers covering a wide range of clinical presentations. Evidence from these studies concerning clinical and cost effectiveness, signposting to appropriate services, patient compliance, and safety was found to be weak or inconsistent [16] and most notably, not conducted in a setting relevant to the UK health and social care system. The methodological challenges documented by other digital intervention researchers could, in part, contribute to the validity of results (29) and we sought to draw on their experience to improve the validity of our main trial results by testing our system in a real-world setting. A randomized controlled trial (RCT) is considered the gold standard methodological design to reduce conscious or unconscious bias, using randomization and blinding to ensure no false conclusions are drawn from the study research [301], with piloting required to ensure trial success. For our study we chose a non-inferiority design, not determining if triage performed using DART was superior to physiotherapist-led triage, rather if it was not "unacceptably worse" (31). This allowed consideration of potential non-clinical benefits such as patient convenience, satisfaction and cost effectiveness. The pilot study described in this paper was to ensure successful delivery of a main trial examining DART safety and efficacy and to assess the suitability of the RCT design for evaluating a digital triage system.

**DART Overview**

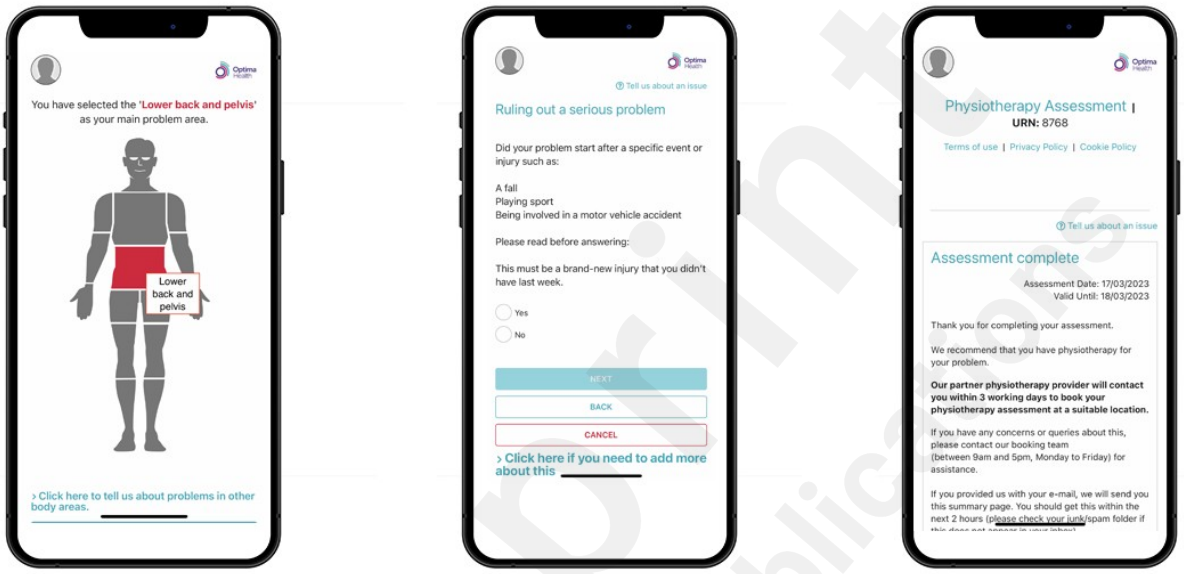
DART (developed by Optima Health) is a web-based first contact mHealth system designed specifically to direct MSK patients to the correct level of care (Figure 1). DART contains an algorithm driving question and response options leading to a triage recommendation configured to match the provider's clinical services, based on evidence-based practice, clinical guidelines, and sector-specific referral criteria. For this reason, there may be variants of DART, containing subtle differences to ensure the algorithm is mapped to the MSK service in which it sits. Triage recommendation options may include emergency or routine medical assessment, physiotherapy, self-management programs or psychological support services. For this study, the DART algorithm was mapped to the specific NHS MSK service delivered at the trial site. DART is a web application, only accessible by users via the MSK service provider's website. Is not intended for general population use via the App store. DART is classified as a "symptom checker" by the UK Medicines and Healthcare products Regulatory Agency (MRHA) and so does not qualify as a Medical Device [25]. It is classified as a tier C system by the UK National Institute of Health and Care Excellence (NICE) whose classification groups align with those proposed by the International Medical Device Regulators Forum [20]. It has been used within a controlled real-world occupational health setting within Optima Health since 2019 with over 9,000 assessments being completed.

**Previous work**

Previous work as described in the pilot protocol [32] included assessment of clinical validity by

an expert panel, real-world usability testing [33] and assessment within a controlled clinical environment.

Figure 1. DART mHealth system user display examples




## Aims and Objectives

In this pilot trial, the research aim was to evaluate trial design, assess procedures and collect exploratory data to assess feasibility of delivering an adequately powered, definitive crossover non-inferiority randomized trial, assessing DART safety and efficacy in an NHS primary care setting.

The primary objective of the trial was to collect and synthesize data (agreement of triage outcome made between DART and physiotherapist-led triage) to define a non-inferiority margin and subsequent sample size calculation for an adequately powered main trial using the principles described by Bujang & Baharum [34]. Agreement was defined as the physiotherapist selecting the same triage recommendation as given by DART (Figure 2).

Secondary process objectives were as follows with associated pre-defined outcomes:

1. Recruitment (recruitment rate targets – 50%, retention – 95%, dropouts- <4).
2. Randomization (equal numbers allocated to each intervention arm, occurrences of allocation concealment failure, introduction of bias).
3. Effectiveness of process implementation (occurrence of non-adherence to study protocol, DART login errors, DART system failures).
4. Burden on patients and clinician (measurement of treatment delays and additional time

requirements, feedback from physiotherapist and researchers concerning trial procedure complexity).
5. Participant satisfaction with using DART (amalgamated System Usability Scale scores), with expectation that a mean score of 80 or more would be achieved, a standard consistent with the previously published DART usability study [33].

## Methods

### Study Design

This 8-week crossover non-inferiority randomized controlled pilot trial was conducted within an NHS primary care setting, using equal randomization of 1:1. The study was designed in accordance with the CONSORT guidelines for pilot and feasibility trial [35], CONSORT guidelines for equivalence and non-inferiority randomized trials [36] and E-HEALTH checklist [37]. Whilst the terms feasibility and pilot are often used interchangeably, the term "pilot" trial was chosen by the authors to reflect that the methodology used would be reproduced in a future definitive randomized controlled trial [35].

All participants underwent an on-line DART assessment and a face-to-face assessment with the on-site physiotherapist. The physiotherapist assessment was intended to reproduce the type of questioning a triage physiotherapist would deliver remotely over the telephone, providing whilst a source of "ground truth" with which to compare the DART outcome, in fact potentially providing greater rigor by virtue of the physiotherapist being able to observe and interact with the patient., The physiotherapist assessment consisted of patient history taking and discussion of symptoms but did not include a physical examination. Only the triage outcome from this element was used for study comparison.

An Integrated Knowledge Translation approach as described by Smith et al [38] was adopted, where the MSK services leader, lead primary care physician and study physiotherapist all helped to shape the research, with the aim of improving its utility and impact. This included discussions of triage routing to improve alignment of the DART algorithm alignment with that of the existing clinical service prior to commencing the pilot. A minimum sample size of 76 participants was chosen based on the estimated stepped rules of thumb from Whitehead et al. [39] to demonstrate an extra small, standardized effect size($\sigma < 0.1$) at a 90% powered main trial.

### Trial setting

The Haydock Medical Centre is a well-established multi-disciplinary primary care practice in the Northwest of England, with 50 staff and clinicians, serving over 15,000 patients. Through links with Health Education Northwest, Manchester and Edge Hill Universities, it provides training for primary care physicians, medical students, nurses, and health care assistants. The more recent introduction in the UK of musculoskeletal First Contact Practitioner (FCP) physiotherapists located within primary care clinics is seen as providing an effective alternative to primary care physician or general practitioner (GP) assessment for musculoskeletal conditions, so potentially freeing up physician appointments [40]. The FCP physiotherapist who participated in this trial is based two days per week at the Centre and patients presenting with musculoskeletal symptoms are either booked directly into the FCP physiotherapist diary instead of seeing a primary care physician or are referred to the FCP physiotherapist by another clinician at the practice. By virtue of enhanced clinical skills beyond that of most triage physiotherapists, an FCP physiotherapist is trained to manage more complex cases and may facilitate diagnostic investigation and refer on to specialist services. For this reason, the on-site FCP physiotherapist was chosen to provide the subjective physiotherapist assessment, to act as

a rigorous study comparator with which to evaluate DART.

**Recruitment**

DART has been designed to triage patients self-referring into primary care for any suspected MSK condition, and therefore is not limited to any specific type or stage of injury. Patients may also be directed by their primary care physician to use DART to confirm the type of MSK care required. Posters and leaflets advertising the study were placed in the practice waiting room. Patients with an MSK condition wishing to access support from the practice (either primary care physician or FCP), were offered the opportunity to participate in the study by the reception team at the point of requesting an appointment, either in person at the practice or by telephone. Patients were given a brief eligibility screen and a short description of the study, with those wishing to participate subsequently booked into a 45-minute slot in the trial diary. This appointment duration allowed both assessments to take place in addition to obtaining informed consent, randomization, and blinding processes.

**Inclusion and Exclusion Criteria**

The study participant inclusion criteria were as follows: (1) adults aged > 18 years; (2) able to speak and read English; (3) registered patient at the primary care practice; (4) current musculoskeletal condition for which they were seeking treatment; (5) able to access the internet either themselves or with the help of family or friend.

The study participant exclusion criteria were as follows: (1) significant physical or cognitive impairments sufficient to limit their ability to follow study-related procedures; (2) unwillingness to follow protocol-related procedures; (3) an existing diagnosis for their condition given by a medical professional within the last 7 days; (4) Optima Health employees. Participants were not paid to participate in the trial.

**Ethical Considerations**

This study received human subjects research ethics approval from the Health Research Authority, London-Surrey Borders Research Ethics Committee on March 24th, 2022 (22/LO/0129).

To support informed consent, on arrival participants were given the Participant Information Sheet (Multimedia Appendix 1) outlining the purpose of the trial and the nature of their participation. This included information about the format of the interaction, potential risks, confidentiality and protection of their personal data, use of their data for analysis (including secondary analysis by expert panel review), anonymity of study findings, and their right to withdraw at any time without prejudice. In addition, this document signposted patients to the Queen Mary University of London Privacy Notice (Multimedia Appendix 2). Patients were made aware no renumeration was to be given for participation. They were then given the opportunity to raise questions with the researcher during the formal consenting process, which was conducted in an allocated treatment room. Formal informed consent was obtained by the researcher and documented using an on-line form (Multimedia Appendix 3). Failure to provide consent resulted in the patient immediately receiving a usual care assessment from the on-site physiotherapist, as per the trial protocol [32].

**Data Collection**

The sequence of assessments was determined by randomization to account for order effects in the crossover design and achieved by block randomization with permuted blocks of random size and without stratification factors to avoid selection bias and unequal arms [41-43]. After gaining consent the researcher used Sealed Envelope software [44] to generate a randomization sequence with a 1:1 allocation ratio between the study arms. Triage outcomes

within DART were matched to those available to the physiotherapist based on usual care approaches in musculoskeletal clinical practice. These outcomes were classified into four categories, Medical care, FCP (referral for assessment with an FCP), Physiotherapy care or Remote self-management. There were further sub-outcomes within each category including levels of urgency (Figure 2) allowing direct comparison between the two types of triage assessment outcome. Levels of care were determined by the clinician's skill set and their access to diagnostic or treatment facilities, with Medical care able to provide the greatest level of support and Remote self-management the least.

Figure 2. All possible triage outcomes and sub-outcomes from DART and physiotherapist-led triage assessment.

| **Medical care** | **First Contact Practitioner (FCP) Physiotherapist** | **Physiotherapy care** | **Remote self-management** |
|---|---|---|---|
| Emergency care (Emergency Department referral) | | Post fracture or surgery physiotherapy | Supported self-management |
| Urgent primary care physician (GP) | Urgent FCP | Physiotherapy referral | On-line support material |
| Routine primary care physician (GP) | Routine FCP | Physiotherapy referral plus psychosocial support | |
| Consultant review | | | |

The physiotherapist assessment was completed within a 20-minute appointment which included standardization of time for each study arm to support blinding. To minimize potential bias, the physiotherapist did not discuss any possible diagnosis or give condition management advice to the participant until their assessment had been completed and their study outcome documented. The DART on-line assessment was completed in a clinic room adjacent to the physiotherapist's room, either before or after the appointment with the physiotherapist, depending on the randomization allocation. The researcher logged participants onto DART using a tablet device and explained they would not be able to assist or discuss any of the questions with them. If the participant said they normally used the internet with help from a family member or friend as a surrogate-seeker [45], the researcher would assist in this way to

navigate through the DART assessment and read text but would not discuss clinical details at any stage with the participant. The participant followed the instructions given by DART until they arrived at the final page where the DART outcome was not visible to either participant or researcher but stored in DART for later retrieval and analysis. Thereafter, the participant completed the System Usability Scale (SUS) questionnaire on-line to measure user satisfaction with DART [46]. Both assessments were completed at the same visit and within 10 minutes of each other to reduce variation in clinical presentation. Once the participant had finished both study assessments and data collection complete, they returned to the physiotherapist, who performed a physical assessment and continued with normal care. Participants could opt out of the study at any point, which did not affect their usual physiotherapist-led management. This process was supported through documented guidance and training delivered by the principal investigator. Blinding was ensured at three different points; 1) the physiotherapist was blinded to group allocation and DART assessment outcomes, 2) participants were blinded to the DART assessment outcome and the physiotherapist triage outcome until they have completed both assessments and the SUS and 3) the analysis and interpretation of the study results was completed by researchers blinded to the intervention group allocation.

**Data analysis**

An independent panel consisting of 3 experts qualified to consultant level in musculoskeletal physiotherapy and general practice, provided consensus on all disagreements between DART and physiotherapy-led triage that would yield a safety concern, which were:

1. When the DART outcome was physiotherapy or self-management, and the physiotherapist outcome was emergency or urgent medical care (emergency department referral or urgent primary care physician)
2. When the DART outcome was self-management and the physiotherapist outcome was physiotherapy, FCP or medical care
3. When the DART outcome was routine care, and the physiotherapist outcome was emergency or urgent care

In addition, a random sample of 10% of the remaining cases that did not yield a safety concern were assessed by the panel to decide what they considered to be the correct outcome. This was based on the participant's presentation from the physiotherapist's assessment clinical record and the DART assessment summary which provided the questions asked and the participant's responses. Where two or more panel members disagreed with the physiotherapist's decision, the panel recommendation was used to provide the definitive outcome against which the DART outcome was compared. In all cases where DART did not agree with the physiotherapist outcome further analysis was performed to ascertain the direction and extent of escalation. Under-escalations (where DART recommended a lower level of clinical support than the physiotherapist) and over-escalations (where DART recommended a higher level of clinical support) were assigned to Level 1,2 or 3, dependant on the difference in number of levels between physiotherapist and DART outcomes. Data collected from DART, physiotherapist, on-line SUS questionnaire and researcher log were entered on to an Excel spreadsheet by the Principal Investigator and checked for accuracy by a second researcher prior to analysis. Qualitative data from the physiotherapist, NHS service lead and researchers regarding the study process was noted during informal post-study debrief discussion sessions.

**Statistical analysis**

The primary analysis was an absolute agreement intraclass correlation coefficient ICC(A,2) estimate with 95% confident intervals between DART and the physiotherapist across all triage outcomes, with a sub-analysis of categories (Medical referral, FCP, Physiotherapy, and Self-management) and adverse triage outcomes. This was calculated using SPSS statistical package software version 23(SPSS Inc, Chicago, IL) and based on a single rating, 2-way random-effects model [47,48]. The ICC was reported with a 95% confidence interval which gave a measure of reliability as described by Koo & Li [47]. This measure of agreement would inform a consensus for the non-inferiority margin required for the main study, in turn facilitating a definitive trial sample size calculation. DART user satisfaction scores were reported as a mean SUS mean score and adjective rating across all participants.

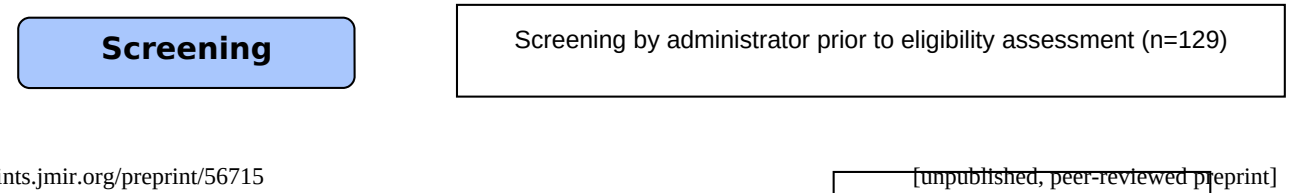
## Results

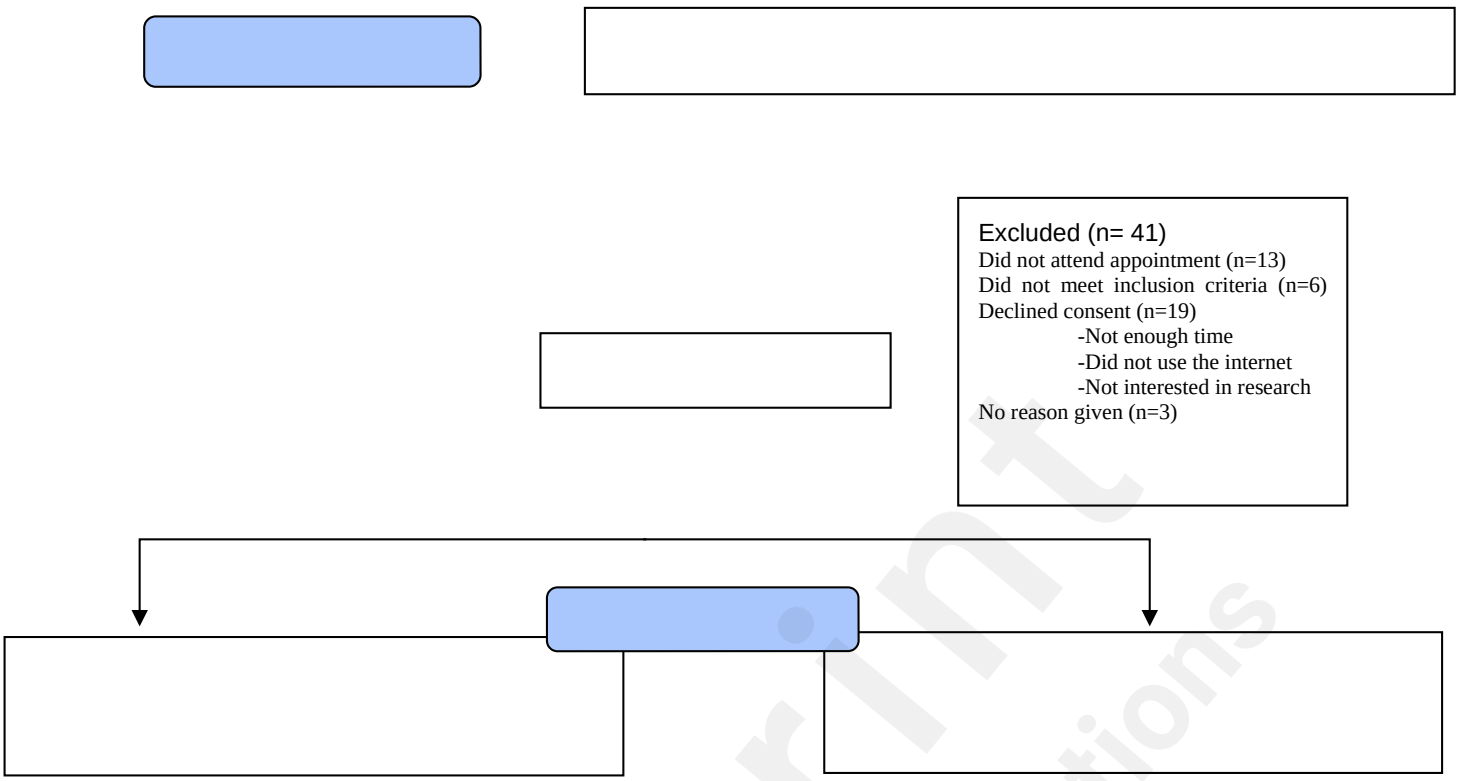
### Recruitment

A total of 129 patients contacted the practice seeking an appointment for a suspected MSK condition during the 8-week trial period, with 92% (119/129) passing initial eligibility screening and being booked into the study (Figure 3). Of these, 35% (41/119) were excluded by the researcher, the main reasons being: 13 not attending their appointment, 6 not meeting the inclusion criteria, and 19 declining consent (with the most common reasons given as not having enough time, did not use the internet and not interested in research). A further 3 patients were unable to participate due to technical issues related to internet connectivity. Recruitment continued until the pre-defined sample size of 76 had been exceeded. A total of 60% (78/129) participants were enrolled in the study. This exceeded the pre-defined recruitment rate of 50%. There were no dropouts and 100% retention of participants, exceeding the pre-defined level of 95%. All data were collected during the single appointment, with no missing data.



*Figure 3. CONSORT diagram showing flow of participants through the study.*






**Randomization and blinding**

Randomization was effective with participants evenly distributed across the two intervention
location concea
x at birth and
successful randomisation and minimised risk of selection bias (Table 2). Bias was minimized through standard timings for both types of assessment, however this meant patients arriving more than 10 minutes late for were unable to participate, with 3
esearchers note
whilst waiting fo
room except when performing study-related activities to minimize this risk.

**Data collection**

Process implementation was effective with full adherence to the trial protocol, evidence of which was documented for each participant on the researcher log. There were 3 patients who would have been eligible to take part, however lack of internet connection within the study area meant they could not participate. Otherwise, there were no DART login errors or DART system failures during data collection. The burden on patients and clinician was considered as acceptable, as there were no treatment delays beyond the 15 minutes taken to complete the study process and participants had no extra travel in addition to that required for their physiotherapy appointment. There was no harm to any participants or unintended effects or consequences. Researchers said the data collection was procedurally complex for them to deliver, particularly around the accuracy of timings to maintain blinding, however the physiotherapist reported their part in the process was straightforward. The additional diary time allocated to the physiotherapist to complete the trial process, over and above usual care was 20 minutes per participant, and for the researcher 45 minutes per participant, the cost of which would need to be factored into the delivery of a future definitive trial.

## Protocol deviations

During the trial the study physiotherapist identified challenges in making decisions for the FCP primary outcome due to ambiguity of the FCP referral definition within the protocol. After discussion between the Principal Investigator and physiotherapist, it was decided to continue as per the study protocol, but once data collection was complete the physiotherapist would review all 22 cases previously routed to FCP and either confirm or amend their outcome prior to data analysis based on clarification of the FCP referral criteria. Demographic characteristics for participants are presented in Table 2. More females were recruited than males (60%: 40%), a ratio higher than reported UK MSK prevalence [7]. The mean age of all participants was 52.9 years (SD 16.79) with a range of 18 to 78 years.

As shown in Figure 4, the most frequently seen age group was 46-65 years old (40%, 31/78,), with the 65+ group representing 27% (21/78). The prevalence of MSK conditions is reported as increasing with age [7], so it is likely that older participants are under-represented in this study. The most frequently selected body sites were low back and pelvis, 23% (18/78), and knee, 22% (17/78) consistent with a recent study examining MSK presentations within a similar urban community primary care practice [49]. Hip conditions represented the next most frequently selected site with 15% (12/78). These three body sites accounted for 60% (47/78) of all presentations.

*Figure 4. Baseline participant demographic characteristics for all participants (n=78) by body site, randomized intervention allocation, sex at birth and age.*

| Body site selected, n=78 | | Head and neck | | Chest and upper back | | Low back and pelvis | | Shoulder | | Elbow | | Wrist and hand | | Hip | | Knee | | Ankle and foot | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | | 5 | | 18 | | 10 | | 3 | | 3 | | 12 | | 17 | | 8 | |
| Intervention allocation A, n=39, B, n=39 | | A | B | A | B | A | B | A | B | A | B | A | B | A | B | A | B | A | B |
| | 18-30 | | | | | ♂♀ | | | | ♂♂ | | ♀ | | ♂♀ | ♂♀♀ | | | ♀ | ♀ |
| Age Mean (SD,range) | 31-40 | | | ♂ | | ♂ | | ♀ | ♂ | | ♂ | | | | | ♀ | ♂ | | ♂ |
| Sex at birth: | 41-50 | | | ♂♂ | | ♀ | ♀♀♀ | ♀♀ | ♂ | | | | | | ♀ | | ♂ | | ♂ |
| ♀ Female 52.66(16.6,19-86) | 51-60 | | | | ♂ | ♀♀ | ♀ | ♂ | | | | | ♂ | | ♀♀ | ♂♀♀♀ | ♂♀♀ | | ♀♀ |
| ♂ Male 53.3(17.35,18-76) | 61-70 | ♀ | ♀ | | | | ♂♂♀♀ | | ♀ | | | ♂ | | ♀♀ | ♀ | ♀ | ♂ | ♂♀ | |
| | 70+ | | | ♀ | | ♂♀♀ | ♀ | ♂ | ♂ | | | | | ♀ | | ♂♂ | ♂♀♀ | | |
| | | | | | *Intervention A participants age* | | | *53.74 (18.07,18-82)* | | | *Intervention B participants age* | | | *52.08 (15.62,19-74)* | | | | | |

## Panel review

The panel of 3 experts reviewed 14 cases (Figure 5). The protocol requirement for a random sample of 10% of participants in addition to the safety cases was exceeded by one case due to researcher error. There was complete agreement between all panel members for 57% (8/14) of cases , and partial agreement between 2 panel members for the remaining 43% (6/14) cases. As per the study protocol, where 2 panel members agreed on the same outcome that differed to that of the physiotherapist, the panel outcome was used for data analysis. This resulted in 3 changes to the physiotherapist's outcome, all of which were the same as the DART outcome. There were 5 cases where 1 panel member disagreed with the physiotherapist's outcome but was insufficient to trigger a change.

Figure 5. Primary outcome by DART, study physiotherapist and expert panel members. DART outcomes classified as adverse (n=5) are shown in red. The remainder represents a further randomly selected 10% sample (n=9)of participants, plus one additional selected in error. Highlighted physiotherapist outcomes signify a panel change, with the revised outcome shown in brackets. Revised

outcomes were used for data analysis (n=3).

| DART | Physiotherapist | Expert 1 | Expert 2 | Expert 3 |
|---|---|---|---|---|
| Self-management | Physiotherapy | Physiotherapy | Physiotherapy | Physiotherapy |
| Self-management | Self-management | Self-management | Self-management | Self-management |
| Self-management | FCP (Self-management) | Self-management | Self-management | Self-management |
| Medical | FCP(Medical) | Medical | Medical | Medical |
| Physiotherapy | FCP | FCP | FCP | Physiotherapy |
| Self-management | Self-management | Self-management | Self-management | Self-management |
| Self-management | Self-management | Self-management | Self-management | Self-management |
| Self-management | FCP | FCP | FCP | FCP |
| Physiotherapy | Physiotherapy | Physiotherapy | Physiotherapy | Physiotherapy |
| Physiotherapy | Physiotherapy | Physiotherapy | Physiotherapy | FCP |
| Physiotherapy | FCP (Physiotherapy) | FCP | Physiotherapy | Physiotherapy |
| Self-management | Physiotherapy | Physiotherapy | Self-management | Physiotherapy |
| Self-management | Physiotherapy | Physiotherapy | Physiotherapy | Self-management |
| Medical | FCP | FCP | FCP | Medical |

The updated physiotherapist outcomes were used in the primary outcome data analysis (Figure 6).

*Figure 6. Physiotherapist and DART primary and secondary outcomes for all participants. Adverse outcomes are shaded.*

| Agree/escalation [Primary outcome] | Escalation Level | Physio primary outcome | DART primary outcome | Physio secondary outcome | DART secondary outcome | Number of cases |
|---|---|---|---|---|---|---|
| Agree | N/A | Medical | Medical | Routine primary care physician [GP] | Routine primary care physician [GP] | 1 |
| Agree | N/A | FCP | FCP | Routine FCP | Routine FCP | 1 |
| Agree | N/A | FCP | FCP | Routine FCP | Urgent FCP | 1 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy referral | Physiotherapy referral | 14 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy +psychosocial support | Physiotherapy referral | 2 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy referral | Physiotherapy + psychosocial support | 1 |
| Agree | N/A | Self management | Self management | Supported self management | Supported self management | 2 |
| Agree | N/A | Self management | Self management | Supported self management | Online support material | 2 |
| Agree | N/A | Self management | Self management | Online support material | Online support material | 7 |
| Agree | N/A | Self management | Self management | Online support material | Supported self management | 2 |
| Under-escalation | Level 1 | FCP | Physiotherapy | Routine FCP | Physiotherapy referral | 3 |
| Under-escalation | Level 1 | Physiotherapy | Self management | Physiotherapy referral | Supported self management | 1 |
| Under-escalation | Level 1 | Physiotherapy | Self management | Physiotherapy referral | Online support material | 3 |
| Under-escalation | Level 2 | Medical | Physiotherapy | Routine primary care physician [GP] | Physiotherapy referral | 2 |
| Under-escalation | Level 2 | FCP | Self management | Routine FCP | Online support material | 1 |
| Over-escalation | Level 1 | FCP | Medical | Urgent FCP | Emergency care | 2 |
| Over-escalation | Level 1 | FCP | Medical | Routine FCP | Emergency care | 2 |
| Over-escalation | Level 1 | FCP | Medical | Routine FCP | Urgent primary care physician [GP] | 1 |
| Over-escalation | Level 1 | Physiotherapy | FCP | Physiotherapy +psychosocial support | Urgent FCP | 2 |
| Over-escalation | Level 1 | Physiotherapy | FCP | Physiotherapy referral | Routine FCP | 2 |
| Over-escalation | Level 1 | Self management | Physiotherapy | Supported self management | Physiotherapy referral | 7 |
| Over-escalation | Level 1 | Self management | Physiotherapy | Online support material | Physiotherapy +psychosocial support | 1 |
| Over-escalation | Level 1 | Self management | Physiotherapy | Online support material | Physiotherapy referral | 8 |
| Over-escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Emergency care | 4 |
| Over-escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Urgent primary care physician [GP] | 1 |
| Over-escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Routine primary care physician [GP] | 1 |
| Over-escalation | Level 2 | Self management | FCP | Online support material | Urgent FCP | 3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Over-escalation | Level 3 | Self management | Medical | Supported self management | Emergency care | 1 |

**Primary objective**

Following the adjustments made by the expert panel, the agreement between physiotherapist and DART across all participants and all primary outcomes was 33/78, 42% (CI 22 to 45), ICC=0.37 with a 95% confidence interval = 0.16-0.55, indicating that the reliability of DART was poor to moderate [47]. Analysis of cases where there was agreement between the physiotherapist and DART by primary outcome is shown in Figure 7.

*Figure 7. Agreement between physiotherapist and DART by primary outcomes for all participants.*

| Primary stratification outcome | Rate, % (95% CI) |
|---|---|
| Medical | 1/3, 33 (0 to 6) |
| FCP | 2/11, 18 (0 to 7) |
| Physiotherapy | 17/31, 55 (10 to 27) |
| Self-management | 13/33, 39 (7 to 22) |

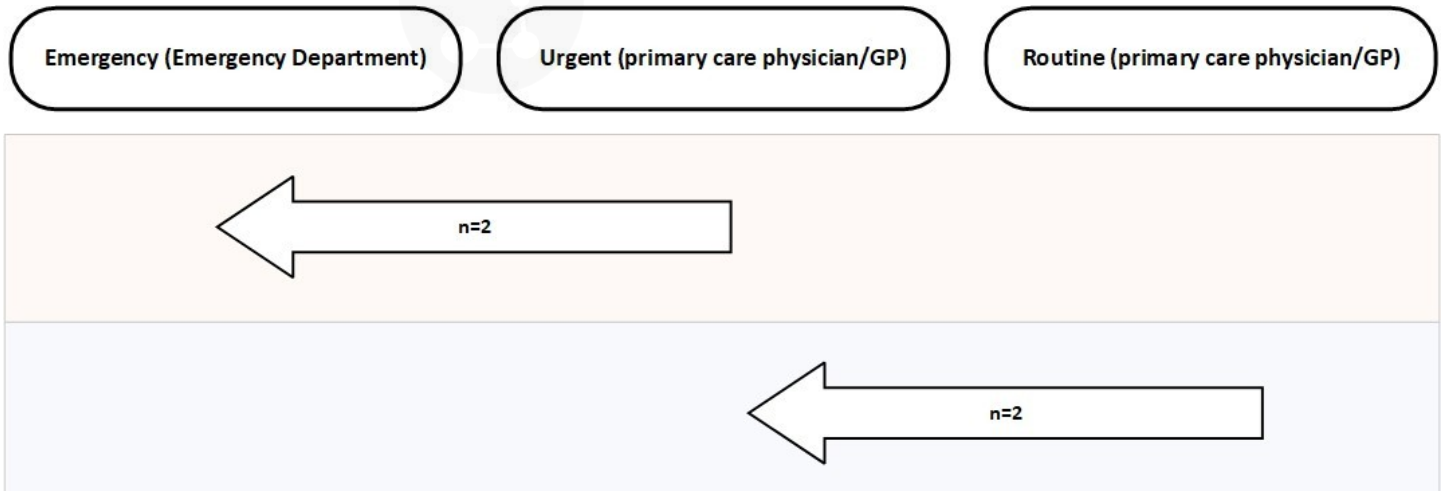
There were just 3 Medical outcomes selected by the physiotherapist, none were emergency or urgent care, all being routine primary physician, DART agreed in 1 case out of the 3. The greatest agreement at 55% was for the Physiotherapy outcome and the lowest was for the FCP outcome, with DART agreeing with only 2 out of the 11 FCP cases presented.

Five cases met the protocol criteria of an adverse outcome representing a potential clinical safety issue: Physiotherapy or Self-management when it should have been Emergency or urgent Medical care (n= 0), Self-management when it should have been either Physiotherapy, FCP or Medical care (n= 5) and routine care when it should have been urgent care (n= 0). In 4 out of 5 of these adverse outcomes DART under-escalated by 1 level to Self-management when the physiotherapist had routed to Physiotherapy. The remaining case was an under-escalation by DART to Self-management when the physiotherapist had routed to FCP. During data analysis it became clear this was created by a foot complaint screening question which has subsequently been revised.

Urgency of referral was defined as secondary outcomes within the Medical and FCP primary outcomes: Emergency (Emergency Department), Urgent (primary care physician (GP) or FCP and Routine (primary care physician (GP) or FCP). Physiotherapy and self-management were considered Routine in terms of urgency. DART over-escalated 6 cases from Routine to Urgent (Figure 8). There were no cases where DART underestimated the urgency of the recommendation, however no participants were deemed by the physiotherapist to require emergency or urgent Medical care, therefore DART routing was not assessed in this area.



*Figure 8. Number of DART over-escalations by urgency of outcome (n=6). DART escalated 6 cases to urgent care where the physiotherapist had recommended routine care. Base of arrow indicates physiotherapist stratification outcome; head of arrow indicates DART outcome.*



1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet [London, England] [Internet]. 2017 Sep 16 [cited 2018 Feb 21];390[10100]:1211–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28919117
2. Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national disability-adjusted life-years [DALYs] for 359 diseases and injuries and healthy life expectancy [HALE] for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392[10159]:1859–922.
3. Sebbag E, Felten R, Sagez F, Sibilia J, Devilliers H, Arnaud L. The world-wide burden of musculoskeletal diseases: A systematic analysis of the World Health Organization Burden of Diseases Database. Ann Rheum Dis. 2019;78[6]:844–8.
4. James SL, Abate D, Abate KH, et al. [2018] Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*; 392: 1789-858. https://doi.org/10.1016/S0140-6736[18]32279-7 PMID: 30496104
5. The Impact of Musculoskeletal Disorders on Americans — Opportunities for Action. Bone and Joint Initiative USA. 2016. [http://www.boneandjointburden.org/docs/BMUSExecutiveSummary2016.pdf].
6. Ahmed N, Ahmed F, Rajeswaran G, Briggs TRW. The NHS must achieve better value from musculoskeletal services. Br J Hosp Med. 2017;78[10]:544–5. PMID: 30548385
7. Versus Arthritis. The State of Musculoskeletal health 2021. https://www.versusarthritis.org/media/24238/state-of-msk-health-2021.pdf
8. Deslauriers S, Déry J, Proulx K, Laliberté M, Desmeules F, Feldman DE, et al. Effects of waiting for outpatient physiotherapy services in persons with musculoskeletal disorders: a systematic review. Disabil Rehabil [Internet]. 2019;0[0]:1–10. Available from: https://doi.org/10.1080/09638288.2019.1639222
9. Lewis AK, Harding KE, Snowdon, David A.1. Lewis AK, Harding KE, Snowdon DA Taylor NF. Reducing wait time from referral to first visit for community outpatient services may contribute to better health outcomes: A systematic review. BMC Health Serv Res. 2018;18[1]:1–14.
10. Williams A, Kamper SJ, Wiggers JH, O'Brien KM, Lee H, Wolfenden L, Yoong SL, Robson E, McAuley JH, Hartvigsen J, Williams CM. Musculoskeletal conditions may increase the risk of chronic disease: a systematic review and meta-analysis of cohort studies. BMC Med. 2018 Sep 25;16[1]:167. doi: 10.1186/s12916-018-1151-2. PMID: 30249247; PMCID: PMC6154805.
11. NHS Improvement [Getting It Right First Time [GIRFT] https://gettingitrightfirsttime.co.uk/what-we-do/#:~:text=Getting%20It%20Right%20First%20Time%20%28GIRFT%29%20is%20a,the%20reduction%20of%20unnecessary%20procedures%2C%20and%20cost%20savings. Accessed 23/01/23
12. C Joseph, C., Morrissey,D., Abdur-Rahman, M., Hussenbux, A., Barton, C. [2014] Musculoskeletal triage: a mixed methods study, integrating systematic review with expert and patient perspectives, *Physiotherapy*, 100[4] 277-89 https://doi.org/10.1016/j.physio.2014.03.007 PMID: 25242531
13. NHS England 75 [2019] https://www.england.nhs.uk/2019/02/support-to-overcome-back-pain/
14. Mallett R., Bakker E., and Burton M. (2014), Is Physiotherapy Self-Referral with Telephone Triage Viable, Cost-effective and Beneficial to Musculoskeletal Outpatients in a

Primary Care Setting?, *Musculoskelet. Care*, 12, 251–260, doi: 10.1002/msc.1075

15. Chartered Society of Physiotherapy. CSP's position on the Home Office 'Shortage Occupation' List. https://www.csp.org.uk/system/files/documents/2019-06/Shortage%20occupations%20jun19.pdf

16. Hill MG,Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. Med J Aust. 2020;212[11]:514–9. PMID: 32391611

17. HM Government. [2020]. Personalised health and care 2020 using data and technology to transform outcomes for patients and citizens. NIB_Report.pdf [publishing.service.gov.uk]

18. Verzantvoort, N. C. M., Teunis, T., Verheij, T. J. M., & van der Velden, A. W. [2018]. Self-triage for acute primary care via a smartphone application: Practical, safe and efficient? In PLoS ONE [Vol. 13, Issue 6]. Public Library of Science. https://doi.org/10.1371/journal.pone.0199284

19. Cowie, J., Calveley, E., Bowers, G., & Bowers, J. [2018]. Evaluation of a digital consultation and self-care advice tool in primary care: A multi-methods study. International Journal of Environmental Research and Public Health, 15[5]. https://doi.org/10.3390/ijerph15050896

20. NICE [2022] 'Evidence standards framework for digital health technologies', NICE, [March], ISBN:978-1-4731-4117-9 https://www.nice.org.uk/corporate/ecd7/resources/evidence-standards-framework-for-digital-health-technologies-pdf-1124017457605

21. Department of Health and Social care [2021] Guidance- A guide to good practice for digital and data-driven health technologies https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology

22. ISO [2019ISO 9241-210:2019 .Ergonomics of Human-System Interaction – Part 210: Human-centred design for interactive systems https://www.iso.org/standard/77520.html

23. European Commission [2014] Green Paper on Mobile Health [mHealth]. SWD135 final. file:///C:/Users/cabella.lowe/AppData/Local/Packages/Microsoft.MicrosoftEdge_8wekyb3d8bbwe/TempState/Downloads/GreenPaperonmobilehealth%20[1].pdf

24. Van Velthoven, M., Smith, J., Wells, G., Brindley, D. [2018] Digital health app development standards: A systematic review protocol. BMJ Open, 8 [8] 1-5 DOI: 10.1136/bmjopen-2018-022969 PMID: 30121614 PMCID: PMC6104800

25. MHRA [2010] Guidance: Medical device stand-alone software including apps [2018]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/717865/Software_flow_chart_Ed_1-05.pdf

26. NHS Digital [2021] NHS digital, data and technology standards framework. Available at: https://digital.nhs.uk/about-nhs-digital/our-work/nhs-digital-data-and-technology-standards/framework

27. NHS Digital [2021] DCB0129:Clinical Risk Management: its Application in the manufacture of Health IT Systems. Available at: https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0129-clinical-risk-management-its-application-in-the-manufacture-of-health-it-systems

**28.** NHS [2021] Digital Technology Assessment Criteria [DTAC]. Available at: https://www.nhsx.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/Bornhöft, L., Larsson, M. E. H., Thorn, J. [2015] Physiotherapy in Primary Care Triage – the effects on utilization of medical services at primary health care clinics by patients and sub-groups of patients with musculoskeletal disorders: a case-control study, *Physiotherapy Theory and Practice - An International Journal of Physical Therapy*, 31[1] doi: 10.3109/09593985.2014.932035. Epub 2014 Jul 2. PMID: 24988315

**29.** Ilicki J [2022] Challenges in evaluating the accuracy of AI-containing digital triage systems: A systematic review. PLoS ONE 17[12]: e0279636. https://doi.org/10.1371/journal.pone.0279636

**30.** The Philosophy of Evidence-Based Medicine, Author[s]:Jeremy Howick,First published:23 March 2011
Print ISBN:9781405196673 |Online ISBN:9781444342673 | DOI:10.1002/9781444342673

**31.** Angeli F, Verdecchia P, Vaudo G, Masnaghetti S, Reboldi G. Optimal Use of the Non-Inferiority Trial Design. Pharmaceut Med. 2020 Jun;34(3):159-165. doi: 10.1007/s40290-020-00334-z. PMID: 32277352.

**32.** Lowe C, Hanuman Sing H, Marsh W, Morrissey D. Validation of a Musculoskeletal Digital Assessment Routing Tool: Protocol for a Pilot Randomized Crossover Noninferiority Trial. JMIR Res Protoc 2021;10[12]:e31541
URL: https://www.researchprotocols.org/2021/12/e31541 DOI: 10.2196/31541

**33.** Lowe C, Browne M, Marsh W, Morrissey D Usability Testing of a Digital Assessment Routing Tool for Musculoskeletal Disorders: Iterative, Convergent Mixed Methods Study J Med Internet Res 2022;24[8]:e38352 doi: 10.2196/38352PMID: 36040787

**34.** Bujang,M., Baharum,N. [2017] Simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Arch Orofac Sci* [2017], **12**[1]: 1-11.

**35.** Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. BMJ. 2016;355. PMID: 27965879

**36.** Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG. Reporting of noninferiority and equivalence randomized trials: Extension of the CONSORT 2010 statement. JAMA - J Am Med Assoc. 2012;308[24]:2594–604. PMID: 16522836

**37.** Eysenbach G. CONSORT-EHEALTH: Implementation of a checklist for authors and editors to improve reporting of Web-based and mobile randomized controlled trials. In: Studies in Health Technology and Informatics. 2013. PMID: 23920638

**38.** Smith,B., Williams., Bone,L. & the Moving Social Work Coproduction Collective [2023] Co-production: A resource to guide co-producing research in the sport, exercise, and health sciences, Qualitative Research in Sport, Exercise and Health, 15:2, 159-187, DOI: 10.1080/2159676X.2022.2052946

**39.** Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res. 2015;25[3]:1057–73. PMID: 26092476

**40.** NHS England [2024] Expanding our workforce, First Contact Physiotherapists https://www.england.nhs.uk/gp/expanding-our-workforce/first-contact-physiotherapists/

**41.** Suresh K. An overview of randomization techniques: An unbiased assessment of outcome in clinical research. Journal of Human Reproductive Sciences. 2011. PMID: 21772732

**42.** Efird J. Blocked randomization with randomly selected block sizes. Int J Environ Res Public Health. 2011;8[1]:15–20. PMID: 21318011

**43.** Broglio K. Randomization in clinical trials: Permuted blocks and stratification. JAMA - J Am Med Assoc. 2018;319[21]:2223–4. PMID: 29872845

**44.** Sealed Envelope Ltd. 2022. Create a blocked randomisation list. [Online] Available from: https://www.sealedenvelope.com/simple-randomiser/v1/lists/1 [Accessed 28 Apr 2023].

**45.** Cutrona, S., Mazor,K., Vieux,S., Luger,T., Volkman,J., Finney Rutten, L. [2015] Health information seeking on behalf of others: Characteristics of 'surrogate seekers'. *J Cancer Educ*. 30[1]: 12–19. doi:10.1007/s13187-014-0701-3.

**46.** Brooke, J. [1996]. SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland [Eds.], Usability Evaluation in Industry. London: Taylor and Francis. IBSN 0-7484-0460Zapata BC, Fernández-Alemán JL, Idri A, Toval A. [2017] 'Empirical Studies on Usability of mHealth Apps : A Systematic Literature Review', [2015]. doi: 10.1007/s10916-014-0182-2. Epub 2015 Jan 20.
PMID: 25600193

**47.** Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med [Internet]. 2016;15[2]:155–63. Available from: http://dx.doi.org/10.1016/j.jcm.2016.02.012

**48.** McGraw KO, Wong SP. Forming Inferences about Some Intraclass Correlation Coefficients. Psychol Methods. 1996;1[1]:30–46. DOI:10.1037/1082-989X.1.1.30

**49.** Keavy R. The prevalence of musculoskeletal presentations in general practice: an epidemiological study. Br J Gen Pract. 2020 Jun;70[suppl 1]:bjgp20X711497. doi: 10.3399/bjgp20X711497. PMID: 32554673.

**50.** Bangor,A., Kortum, P., Miller,J. [2009] Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. J Usability Studies, 4[3]

**51.** Jeff Sauro, James R. Lewis, Chapter 8 - Standardized usability questionnaires, Editor[s]: Jeff Sauro, James R. Lewis, Quantifying the User Experience [Second Edition], Morgan Kaufmann, 2016, Pages 185-248, ISBN 9780128023082, https://doi.org/10.1016/B978-0-12-802308-2.00008-4

**52.** Sauro, J. [2012] A Practical Guide to the System usability Scale: Background, Benchmarks & Best Practices. IBSN:978-1461062707

**53.** Bangor,A., Kortum,P., Miller,J.[2008] An empirical evaluation of the system usability scale. Int J Human-Computer Interaction 24[6] 574-594 doi; 10.1080/10447310802205776

**54.** Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986 Feb 8;1[8476]:307-10. PMID: 2868172.

**55.** Jones, D., Dunn, L., Macleod, U., & Watt, I. [2019]. Safety netting for primary care: Evidence from a literature review. British Journal of General Practice, 69[678], E70–E79. https://doi.org/10.3399/bjgp18X700193

**56.** Gilbert S, Mehl A, Baluch A, *et al*. How accurate are digital symptom assessment apps for suggesting
conditions and urgency advice? A clinical vignettes comparison to GPs. *BMJ Open* 2020;10:e040269. doi:10.1136/bmjopen-2020-040269

**57.** van Teijlingen E, Hundley V. The importance of pilot studies. Nurs Stand. 2002 Jun 19-25;16[40]:33-6. doi: 0.7748/ns2002.06.16.40.33.c3214. PMID: 12216297.

**58.** Chambers D, Cantrell AJ, Johnson M, *et al* Digital and online symptom checkers and

health assessment/triage services for urgent health problems: systematic review. *BMJ Open* 2019;**9:**e027743. doi: 10.1136/bmjopen-2018-027743

**59.** Razzaki, S., Baker,A., Perov, Y., Middleton, K., Baxter, J., Mullarkey, D., et al. A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis. London: Babylon;2018.

**60.** Entezarjou A, Bonamy AE, Benjaminsson S, Herman P, Midlöv P Human- Versus Machine Learning–Based Triage Using Digitalized Patient Histories in Primary Care: Comparative Study JMIR Med Inform 2020;8[9]:e18930 doi: 10.2196/18930PMID: 32880578PMCID: 7499160

**61.** Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, Normahani P, Ashrafian H, Markar SR, Sounderajah V, Darzi A. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. NPJ Digit Med. 2022 Aug 17;5[1]:118. doi: 10.1038/s41746-022-00667-w. PMID: 35977992; PMCID: PMC9385087.

**62.** Hahn S. [2012] Understanding noninferiority trials. Korean J Pediatr. 2012 Nov;55[11]:403-7. doi: 10.3345/kjp.2012.55.11.403. Epub 2012 Nov 23. PMID: 23227058; PMCID: PMC3510268.

**63.** Nasa, P., Jain, R., & Juneja, D. [2021]. Delphi methodology in healthcare research: How to decide its appropriateness. World Journal of Methodology, 11[4], 116–129. https://doi.org/10.5662/wjm.v11.i4.116

**64.** Evans SC, Roberts MC, Keeley JW, Blossom JB, Amaro CM, Garcia AM, Stough CO, Canter KS, Robles R, Reed GM. Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. Int J Clin Health Psychol. 2015 May-Aug;15[2]:160-170. doi: 10.1016/j.ijchp.2014.12.001. Epub 2015 Jan 29. PMID: 30487833; PMCID: PMC6224682.

**65.** Semigran, H. L., Linder, J. A., Gidengil, C., & Mehrotra, A. [2015]. Evaluation of symptom checkers for self diagnosis and triage. Medical Journal, 351. https://doi.org/10.2307/26522266

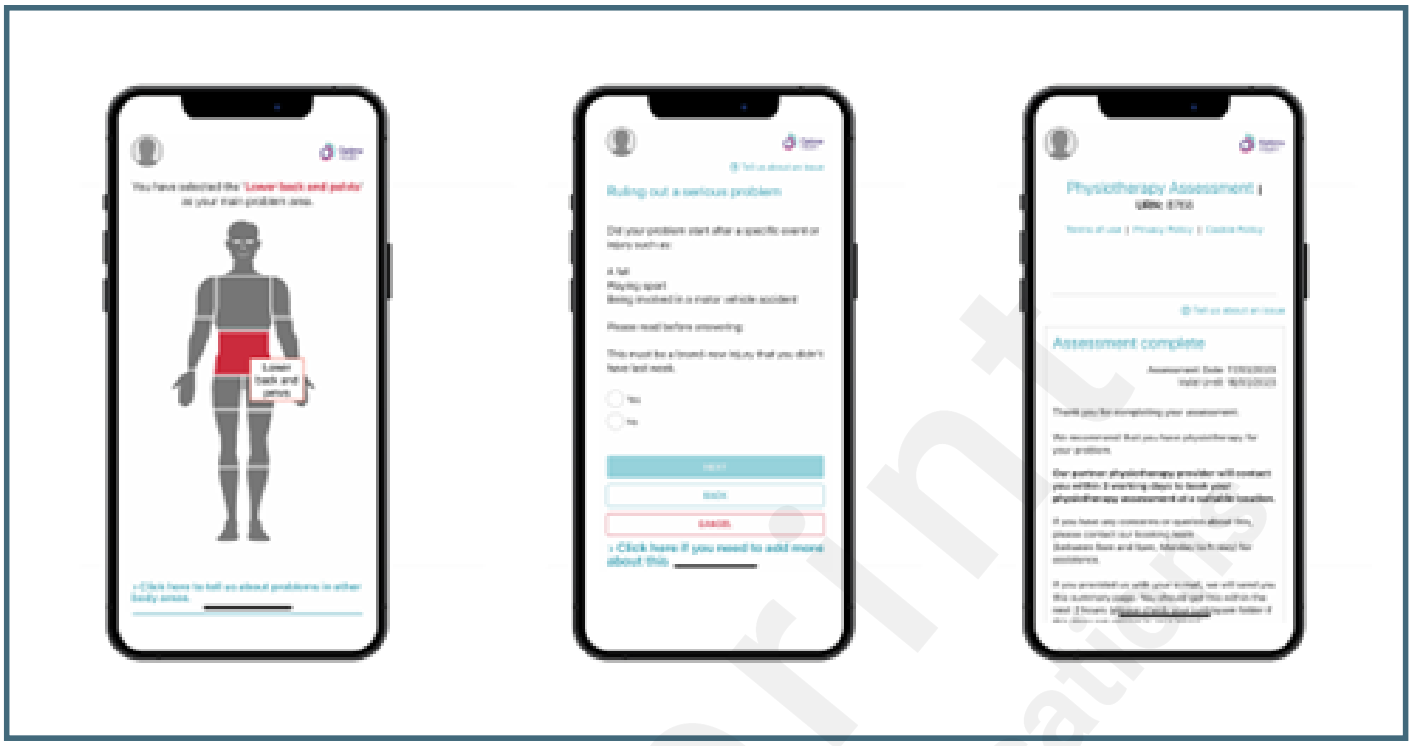
**66.** Leggat, F., R. Wadey, M. Day, S. Winter, and P. Sanders. 2021. “Bridging the Know-do Gap Using Integrated Knowledge Translation and Qualitative Inquiry: A Narrative Review.” Qualitative Research in Sport, Exercise and Health 1–14. doi:10.1080/2159676X.2021.1954074.

**67.** Gottliebsen K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. BMJ Health Care Inform. 2020 May;27[1]:e100114. doi: 10.1136/bmjhci-2019-100114. PMID: 32385041; PMCID: PMC7245402.

# Supplementary Files

# Figures

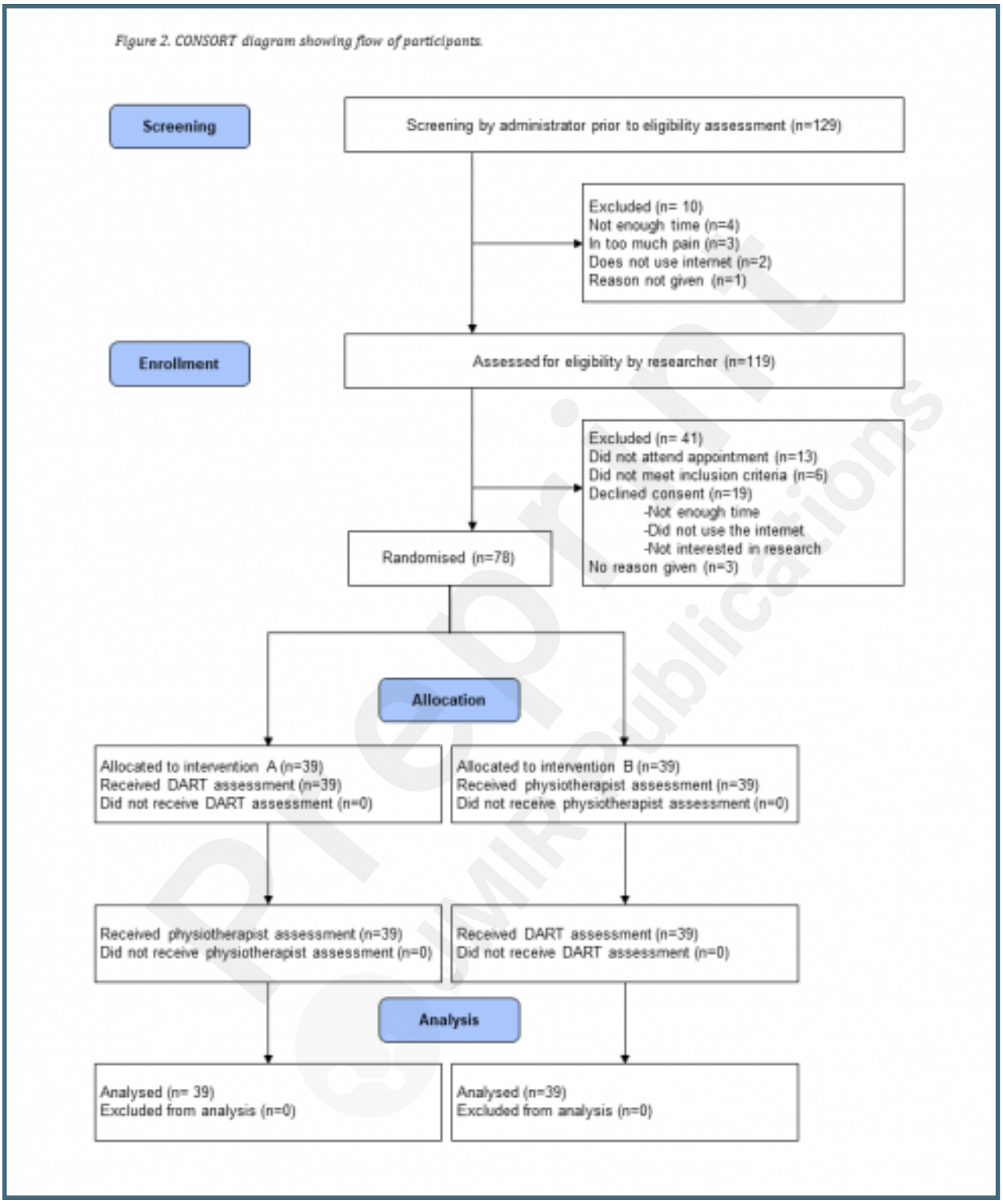
DART mHealth system.



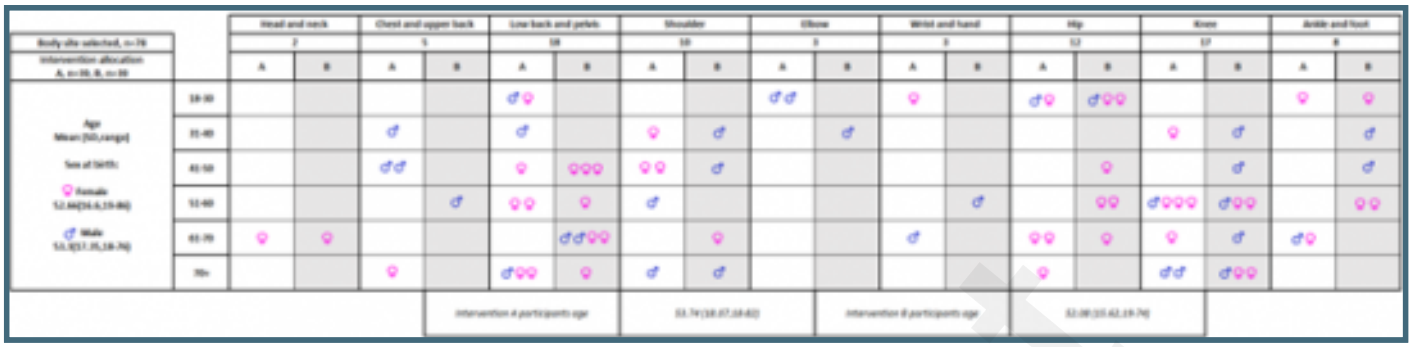
All possible triage outcomes and sub-outcomes from DART and physiotherapist-led triage assessment.

| Medical care | First Contact Practitioner (FCP) Physiotherapist | Physiotherapy care | Remote self-management |
|---|---|---|---|
| **Emergency care (Emergency Department referral)** | | Post fracture or surgery physiotherapy | Supported self-management |
| **Urgent primary care physician (GP)** | Urgent FCP | Physiotherapy referral | On-line support material |
| **Routine primary care physician (GP)** | Routine FCP | Physiotherapy referral plus psychosocial support | |
| **Consultant review** | | | |

CONSORT diagram showing flow of participants.



Baseline demographic characteristics by body site, randomized intervention allocation, sex at birth and age.



Primary outcome by DART, study physiotherapist and expert panel members. DART outcomes classified as adverse (n=5) are shown in red. The remainder represents a further randomly selected 10% sample (n=9)of the remaining cases. Highlighted physiotherapist outcomes signify a panel change, with the revised outcome shown in brackets. Revised outcomes were used for data analysis (n=3).

| Agree/escalation (Primary outcome) | Escalation level | Physio primary outcome | DART primary outcome | Physio secondary outcome | DART secondary outcome | Number of cases |
|---|---|---|---|---|---|---|
| Agree | N/A | Medical | Medical | Routine primary care physician (GP) | Routine primary care physician (GP) | 1 |
| Agree | N/A | FCP | FCP | Routine FCP | Routine FCP | 1 |
| Agree | N/A | FCP | FCP | Routine FCP | Urgent FCP | 1 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy referral | Physiotherapy referral | 18 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy +psychosocial support | Physiotherapy referral | 2 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy referral | Physiotherapy + psychosocial support | 1 |
| Agree | N/A | Self management | Self management | Supported self management | Supported self management | 2 |
| Agree | N/A | Self management | Self management | Supported self management | Online support material | 2 |
| Agree | N/A | Self management | Self management | Online support material | Online support material | 7 |
| Agree | N/A | Self management | Self management | Online support material | Supported self management | 2 |
| Under escalation | Level 1 | FCP | Physiotherapy | Routine FCP | Physiotherapy referral | 3 |
| Under escalation | Level 1 | Physiotherapy | Self management | Physiotherapy referral | Supported self management | 1 |
| Under escalation | Level 1 | Physiotherapy | Self management | Physiotherapy referral | Online support material | 1 |
| Under escalation | Level 2 | Medical | Physiotherapy | Routine primary care physician (GP) | Physiotherapy referral | 2 |
| Under escalation | Level 2 | FCP | Self management | Routine FCP | Online support material | 1 |
| Over escalation | Level 1 | FCP | Medical | Urgent FCP | Emergency care | 2 |
| Over escalation | Level 1 | FCP | Medical | Routine FCP | Emergency care | 2 |
| Over escalation | Level 1 | FCP | Medical | Routine FCP | Urgent primary care physician (GP) | 1 |
| Over escalation | Level 1 | Physiotherapy | FCP | Physiotherapy +psychosocial support | Urgent FCP | 2 |
| Over escalation | Level 1 | Physiotherapy | FCP | Physiotherapy referral | Routine FCP | 2 |
| Over escalation | Level 1 | Self management | Physiotherapy | Supported self management | Physiotherapy referral | 7 |
| Over escalation | Level 1 | Self management | Physiotherapy | Online support material | Physiotherapy +psychosocial support | 1 |
| Over escalation | Level 1 | Self management | Physiotherapy | Online support material | Physiotherapy referral | 8 |
| Over escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Emergency care | 4 |
| Over escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Urgent primary care physician (GP) | 1 |
| Over escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Routine primary care physician (GP) | 1 |
| Over escalation | Level 2 | Self management | FCP | Online support material | Urgent FCP | 3 |
| Over escalation | Level 3 | Self management | Medical | Supported self management | Emergency care | 1 |

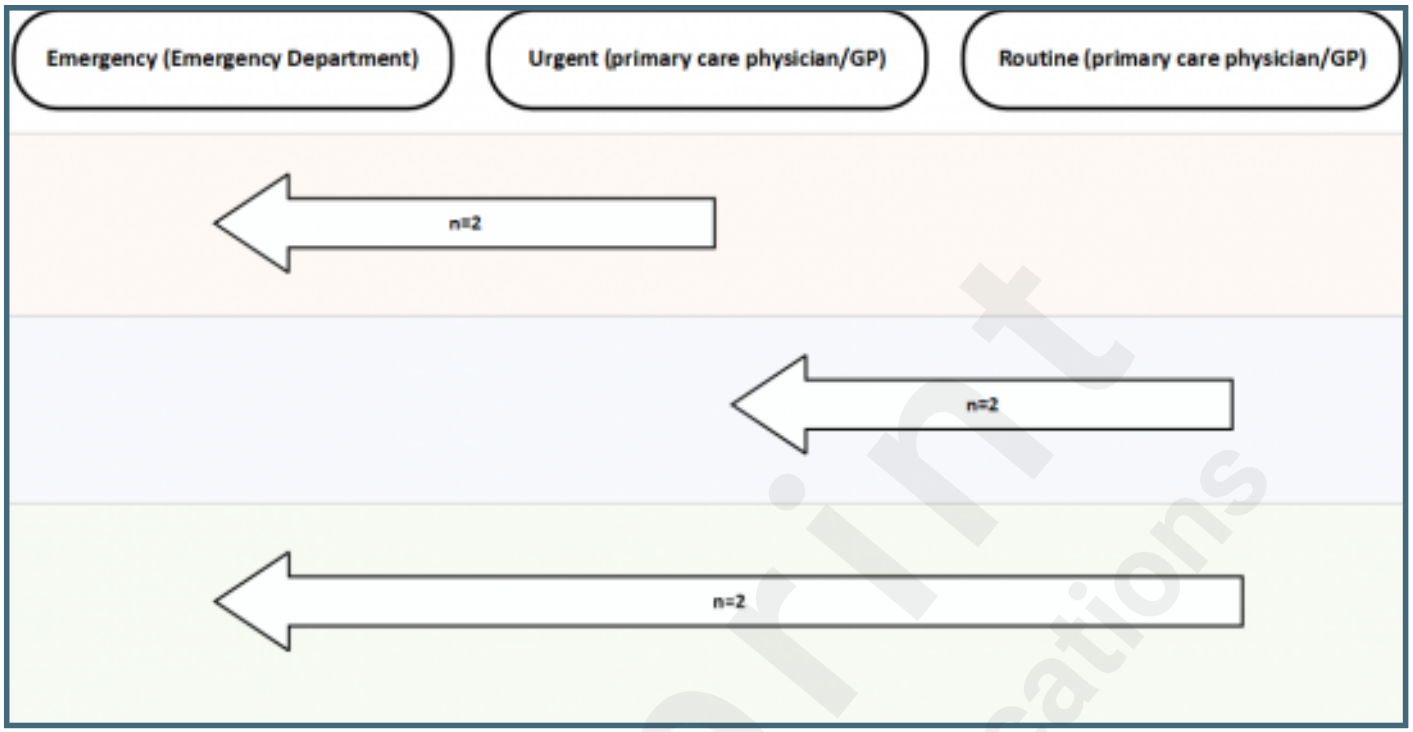
Physiotherapist and DART primary and secondary outcomes for all participants. Adverse outcomes are shaded.

| Agree/escalation (Primary outcome) | Escalation level | Physio primary outcome | DART primary outcome | Physio secondary outcome | DART secondary outcome | Number of cases |
|---|---|---|---|---|---|---|
| Agree | N/A | Medical | Medical | Routine primary care physician (GP) | Routine primary care physician (GP) | 1 |
| Agree | N/A | FCP | FCP | Routine FCP | Routine FCP | 1 |
| Agree | N/A | FCP | FCP | Routine FCP | Urgent FCP | 1 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy referral | Physiotherapy referral | 16 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy +psychosocial support | Physiotherapy referral | 2 |
| Agree | N/A | Physiotherapy | Physiotherapy | Physiotherapy referral | Physiotherapy + psychosocial support | 1 |
| Agree | N/A | Self management | Self management | Supported self management | Supported self management | 2 |
| Agree | N/A | Self management | Self management | Supported self management | Online support material | 2 |
| Agree | N/A | Self management | Self management | Online support material | Online support material | 7 |
| Agree | N/A | Self management | Self management | Online support material | Supported self management | 2 |
| Under-escalation | Level 1 | FCP | Physiotherapy | Routine FCP | Physiotherapy referral | 3 |
| Under-escalation | Level 1 | Physiotherapy | Self management | Physiotherapy referral | Supported self management | 1 |
| Under-escalation | Level 1 | Physiotherapy | Self management | Physiotherapy referral | Online support material | 1 |
| Under-escalation | Level 2 | Medical | Physiotherapy | Routine primary care physician (GP) | Physiotherapy referral | 2 |
| Under-escalation | Level 2 | FCP | Self management | Routine FCP | Online support material | 1 |
| Over-escalation | Level 1 | FCP | Medical | Urgent FCP | Emergency care | 2 |
| Over-escalation | Level 1 | FCP | Medical | Routine FCP | Emergency care | 2 |
| Over-escalation | Level 1 | FCP | Medical | Routine FCP | Urgent primary care physician (GP) | 1 |
| Over-escalation | Level 1 | Physiotherapy | FCP | Physiotherapy +psychosocial support | Urgent FCP | 2 |
| Over-escalation | Level 1 | Physiotherapy | FCP | Physiotherapy referral | Routine FCP | 2 |
| Over-escalation | Level 1 | Self management | Physiotherapy | Supported self management | Physiotherapy referral | 7 |
| Over-escalation | Level 1 | Self management | Physiotherapy | Online support material | Physiotherapy +psychosocial support | 1 |
| Over-escalation | Level 1 | Self management | Physiotherapy | Online support material | Physiotherapy referral | 8 |
| Over-escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Emergency care | 4 |
| Over-escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Urgent primary care physician (GP) | 1 |
| Over-escalation | Level 2 | Physiotherapy | Medical | Physiotherapy referral | Routine primary care physician (GP) | 1 |
| Over-escalation | Level 2 | Self management | FCP | Online support material | Urgent FCP | 3 |
| Over-escalation | Level 3 | Self management | Medical | Supported self management | Emergency care | 1 |

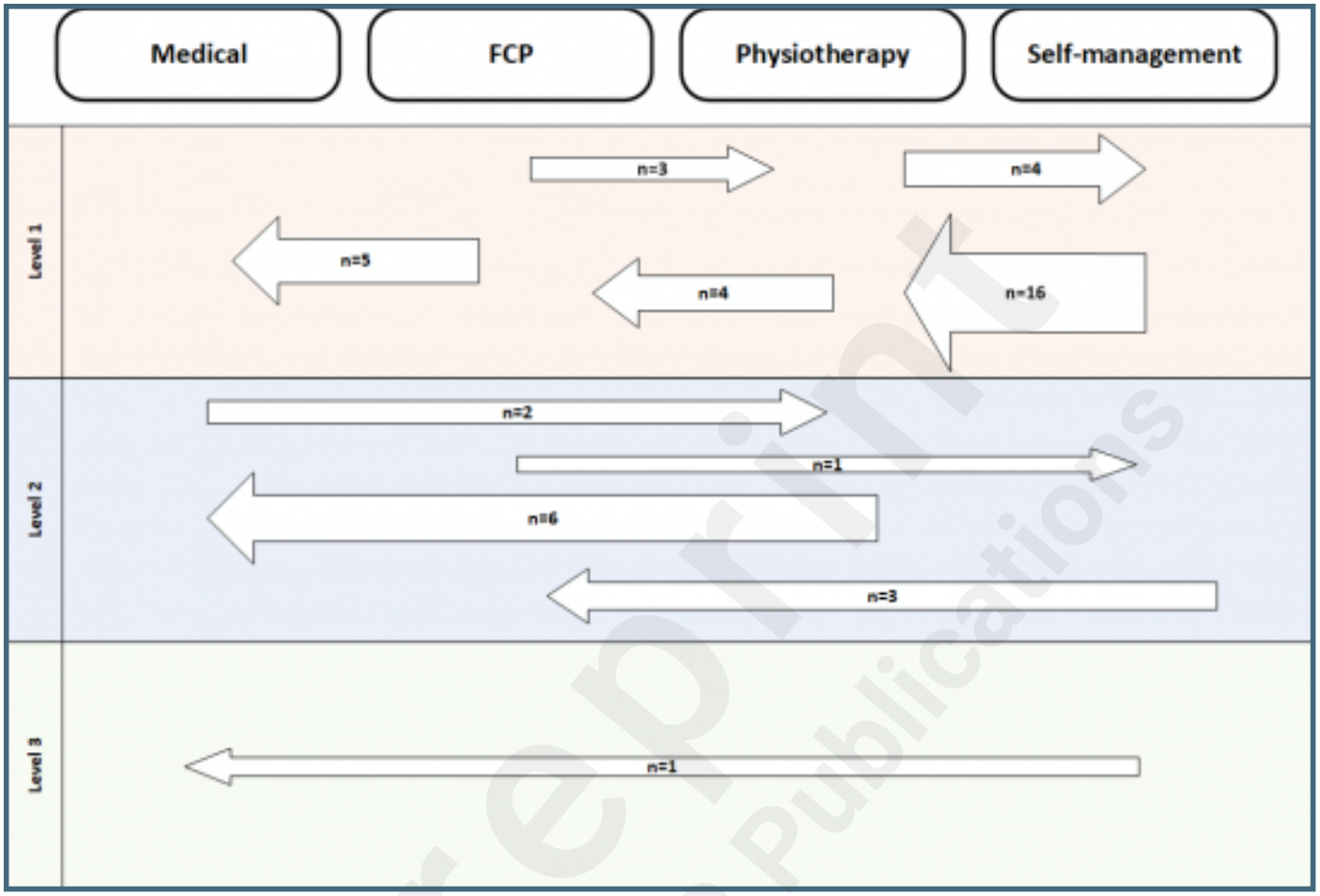
Agreement between physiotherapist and DART by primary outcomes.

| Primary stratification outcome | Rate, % (95% CI) |
|---|---|
| Medical | 1/3, 33 (0 to 6) |
| FCP | 2/11, 18 (0 to 7) |
| Physiotherapy | 17/31, 55 (10 to 27) |
| Self-management | 13/33, 39 (7 to 22) |

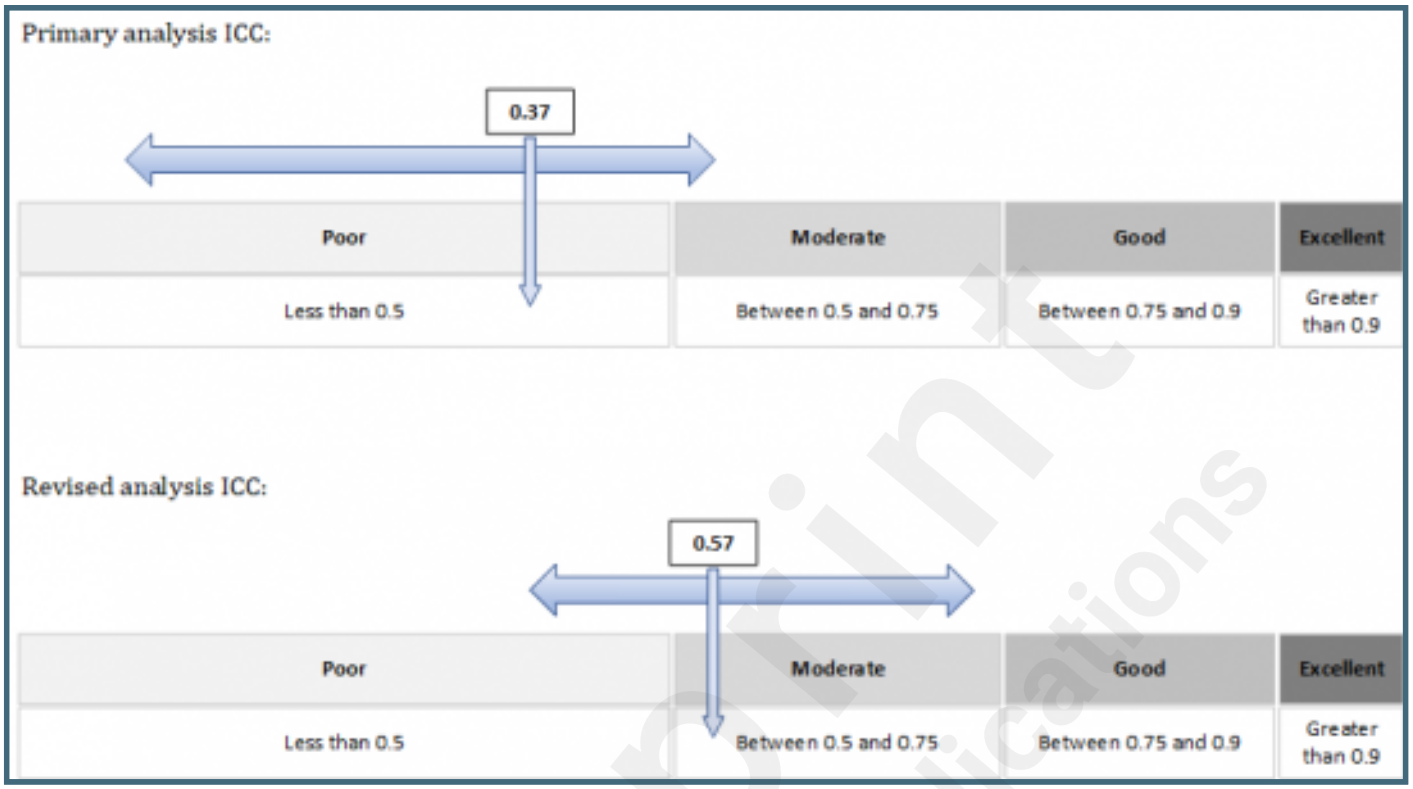
DART escalations by urgency of outcome. Base of arrow indicates physiotherapist stratification outcome; head of arrow indicates DART outcome.



All DART escalations. Arrows pointing to the right indicate under-escalations, those pointing to the left over-escalations. Base of arrow indicates physiotherapist outcome, head of arrow indicated DART outcome. Size of arrows are indicative of volume of cases.



Primary and adjusted analysis calculation of Intraclass Correlation Coefficient and associated confidence interval indicating DART reliability.



# Multimedia Appendixes

Patient Information Sheet.
URL: http://asset.jmir.pub/assets/648a5a6b8bd9fdd8dc593ec396b8b523.docx

Queen Mary University of London participant Privacy Notice.
URL: http://asset.jmir.pub/assets/b82cf2bf9cf1168ab1cbc801e78ff272.pdf

Participant consent form.
URL: http://asset.jmir.pub/assets/63e9a81d1efd6e4ebf44dbed6a98cda8.docx

# CONSORT (or other) checklists

CONSORT EHealth Checklist.
URL: http://asset.jmir.pub/assets/14f4c58208653cf403096924a3829a03.pdf

# Related publication(s) - for reviewers eyes onlies

DART pilot study protocol.
URL: http://asset.jmir.pub/assets/3d185fbf10acd4969790a9d8f90f9a05.pdf

# TOC/Feature image for homepages

Digital Assessment Routing Tool (DART).