Data collection to develop data driven algorithms for predicting the right advice at the right time in patients with hip and knee osteoarthritis: the e-cOAch cross-over study

(May 2025)

May 2015: adaptation section 11.5: text in accordance to old and new Measure regarding Compulsory Insurance for Clinical Research in Humans
Sept 2015: adaptation section 9.1, 9.2 and 12.5: text in accordance to WMO amendment on reporting SAE and temporary halt (section 10 of WMO)
Oct 2015: adaptation section 4.4 – comment [CCMO15], 8.2 and 10.1 with respect to methodology/statistics
Sept 2018: adaptation section 12.1 and comment [CCMO46] due to applicability GDPR as of May, 2018

-

-

-

-

-

PROTOCOL TITLE '**Data collection to develop data driven algorithms for predicting the right advice at the right time in patients with hip and knee osteoarthritis: the e-cOAch cross-over study**

Protocol ID	**KICH1.GZ03.21.005**
Short title	**e-cOAch cross-over study**
EudraCT number	***Not applicable***
Version	*5*
Date	*20-05-2025*
Project leader	Martijn Pisters, PhD, PT Dept. of Rehabilitation, Physiotherapy Science and Sport, UMC Utrecht, Utrecht University; Research Group Empowering Healthy Behaviour, Dept. of Health Innovations & Technology, Fontys University of Applied Sciences; Center for Physiotherapy Research and Innovation in Primary Care, Julius Health Care Centers. m.f.pisters@umcutrecht.nl
Project investigator team	Femke Groen, MSc, PT Dept. of Rehabilitation, Physiotherapy Science and Sport, UMC Utrecht, Utrecht University; f.groen-4@umcutrecht.nl M. de Rooij, PhD, PT Dept. of Rehabilitation, Physiotherapy Science and Sport, UMC Utrecht, Utrecht University; m.derooij-8@umcutrecht.nl
Principal investigator(s) (in Dutch: hoofdonderzoeker/ uitvoerder)	Martijn Pisters, PhD, PT Dept. of Rehabilitation, Physiotherapy Science and Sport, UMC Utrecht, Utrecht University; Research Group Empowering Healthy Behaviour, Dept. of Health Innovations & Technology, Fontys University of Applied Sciences; Center for Physiotherapy Research and Innovation in Primary Care, Julius Health Care Centers. m.f.pisters@umcutrecht.nl
Sponsor (in Dutch: verrichter/opdrachtgever)	University Medical Centre Utrecht

Subsidising party	*NWO*
Independent expert (s)	*F. Hartgens, PhD, PT, MD* *L.M.G.Hartgens@umcutrecht.nl*

PROTOCOL SIGNATURE SHEET

Name	**Signature**	**Date**
Head of Department: ***Prof. Dr. J.M.A. Visser-Meily*** ***Dept. of Rehabilitation, Physiotherapy Science and Sport, University Medical Centre Utrecht*** *j.m.a.visser-meily@umcutrecht.nl* ***Tel. (+31) 88 755 5881***		
Project leader Martijn Pisters, PhD, PT Dept. of Rehabilitation, Physiotherapy Science and Sport, UMC Utrecht, Utrecht University; Research Group Empowering Healthy Behaviour, Dept. of Health Innovations & Technology, Fontys University of Applied Sciences; Center for Physiotherapy Research and Innovation in Primary Care, Julius Health Care Centers. m.f.pisters@umcutrecht.nl		

LIST OF ABBREVIATIONS AND RELEVANT DEFINITIONS

AE	**Adverse Event**
BMI	**Body Mass Index**
CCMO	**Central Committee on Research Involving Human Subjects; in Dutch: Centrale Commissie Mensgebonden Onderzoek**
DSMB	**Data Safety Monitoring Board**
EU	**European Union**
EudraCT	**European drug regulatory affairs Clinical Trials**
GCP	**Good Clinical Practice**
GDPR	**General Data Protection Regulation; in Dutch: Algemene Verordening Gegevensbescherming (AVG)**
IB	**Investigator's Brochure**
METC	**Medical research ethics committee (MREC); in Dutch: medisch-ethische toetsingscommissie (METC)**
OA	**Osteoarthritis**
(S)AE	**(Serious) Adverse Event**
Sponsor	**The sponsor is the party that commissions the organisation or performance of the research, for example a pharmaceutical company, academic hospital, scientific organisation or investigator. A party that provides funding for a study but does not commission it is not regarded as the sponsor, but referred to as a subsidising party.**
SUSAR	**Suspected Unexpected Serious Adverse Reaction**
WMO	**Medical Research Involving Human Subjects Act; in Dutch: Wet Medisch-wetenschappelijk Onderzoek met Mensen**

SUMMARY

Rationale: Guidelines recommend a stepped-care strategy for patients with osteoarthritis (OA), beginning with non-operative approaches. However, these treatments are often underutilized. Artificial intelligence (AI) algorithms have the potential to offer just-in-time guidance, highlighting the need for further development. Understanding data-driven factors that predict OA complaints is crucial for the development of future AI models.

Objective: The objective of this study is to collect data to develop data-driven models to predict changes over time in physical functioning, which can be used within the ArtroseCoach web application to provide just-in-time personalized self-management advice.

Study design: This study is a prospective longitudinal observational crossover study. Participants will be enrolled for one year, with data collection occurring every two weeks. Participants will be randomly assigned to one of the self-care programs within the ArtroseCoach web application (physical activity, weight management, sleep), or no intervention at week 3, 15, 27 and 39. No participant receives the same advice twice.

Study population: People with OA on the hip and/or knee, diagnosed using the self-administered NICE criteria.

Intervention: Participants use the ArtroseCoach 1.0 web application to improve self-management. The content of the ArtroseCoach web application consists of different programs with education about the disease, lifestyle advice and behaviour change support (e.g., physical activity program, weight management program, sleep program).

Main study parameters/endpoints: The main outcomes of the study are changes in pain, physical functioning and participation over time.

Nature and extent of the burden and risks associated with participation, benefit and group relatedness: The risks for participants are expected to be low, since advices follow the current literature and guidelines and are focused on self-management and lifestyle. The burden of the data collection is relatively high, consisting questionnaires every two weeks during 12 months. Additionally, a subset of participants will be asked to wear a wearable on their wrist during 12 months. Participants may benefit, as improvements in lifestyle behaviours could lead to improved physical functioning and participation.

1. INTRODUCTION AND RATIONALE

Background

In the Netherlands, 1.6 million people have osteoarthritis (OA)[1]. The prevalence is expected to increase by nearly 200% by 2050 due to an aging population and increasing obesity numbers[1]. The healthcare costs associated with OA amount to 1.2 billion euros annually and continue to rise[2]. The most often affected joints are the hip and knee. Symptoms include pain and stiffness, often leading to unhelpful beliefs about the disease and insufficient levels of physical activity, which ultimately result in reduced physical functioning and participation in daily life.

To optimize physical functioning and participation, OA guidelines recommend a stepped care strategy. Stepped care recommends that more advanced options (e.g joint replacement) should only be considered if options listed in previous steps failed to produce satisfactory results regarding osteo-arthritis complaints[34]. Step 1 involves self-management education and advice on lifestyle such as movement, sleep and weight management. Step 2 consists of nonsurgical treatment by health care professionals (HCPs) such as (guided) exercise therapy, dietary care, and pain medication, often com-bined with self-management. Lastly, step 3 involves specialized and multidisciplinary care, including surgery. If options in step 1 are not sufficient, steps 2 and 3 follow consecutively[4].

Despite these guidelines, stepped care is insufficiently implemented in clinical practice, resulting in underuse of self-management, exercise, and weight management[5,6], and overuse of surgical interventions[5]. Approximately half of individuals with OA undergo joint replacement without receiving evidence-based nonsurgical treatments first[5]. Step 1 treatments, such as acetaminophen and lifestyle advice, were received by 79% and 60% of patients[5]. Step 2 treatments, including exercise-based therapy and diet therapy, were received by 66% and 19%[5]. Step 3, involving intra-articular injection,was received by 47%. Non-surgical treatment utilization was lower than in 2013[5]. Moreover, many individuals with OA lack the ability to self-manage and control their disease effectively, resulting in suboptimal health decisions[7]. To increase their empowerment during their OA journey, more support in self-management is necessary.

Currently, no suitable and inclusive tools exist to guide people in OA self-management. This lack of guidance, especially in the early stages of OA, result in reduced self-management and physical functioning. Especially since the population with OA is heterogeneous (e.g., age, degree of complaints, BMI, coping, and health literacy skills), and the course of OA also varies, guidance should be better tailored to the needs of the individual at a specific moment: the right advice at the right time. Additionally, it's crucial to recognize that the support needed for effective self-management varies among individuals, encompassing a range of needs from dietary modifications and weight management to increased physical activity and improved sleep quality. Despite the chronic nature of OA, the condition's intermittent nature allows individuals to largely self-manage their journey. Healthcare providers play a supportive role, stepping in as needed to offer guidance and assistance, empowering individuals to take control of their own care. Therefore, optimizing self-management support throughout the OA journey, especially during periods without active healthcare professional involvement, is crucial.

Due to the growing burden of OA on the health system, and the rapid advances in technology, delivering the intervention through digital technologies could be an economical and effective community-based model of care. Lessons from earlier projects are that people with OA like to be coached with the use of digital health[8–10]. The use of continuous data collection in combination with artificial intelligence (AI) has high potential to provide the right advice at the right moment.

To provide just-in-time personalized guidance, knowledge about the prediction of the right time of guidance, as well as the most appropriate content of guidance and behaviour change strategies is needed. However, there is currently a lack of AI algorithms specifically designed for these purposes in OA, emphasizing the need for further development. Current predictive models rely to a large extent on complex measurements such as MRI and X-ray[11,12]. Understanding which self-measured factors predict worsening OA symptoms is essential for developing future AI models. It is also important to consider psychosocial factors that influence OA-related behaviors, such as physical activity, sleep, and weight management. Additionally, for an AI model designed to support long-term self-management, we need to determine how to predict symptom deterioration with as few measurements as possible. Finally, if the timing of AI-generated advice is crucial, the model must be guided by data that identifies the most appropriate moment to provide recommendations.

Therefore, in the present study, the objective is to collect data to develop data-driven models to predict changes over time in pain, physical function and participation, which can be used within the ArtroseCoach web application to provide just-in-time personalized self-management advice. We strive to be inclusive by considering individuals with varying levels of digital health literacy skills.

The present study is part of a 6-year project: SMART (Figure 1). The overarching aim is to develop and test an AI-based app which provides just-in-time advice and suggestions, tailored to the distinctive characteristics of users, through suitable behaviour change strategies and in line with the guidance/ information needed in according to existing OA guidelines. The present METC application describes part 2a of the observational cross-over phase.

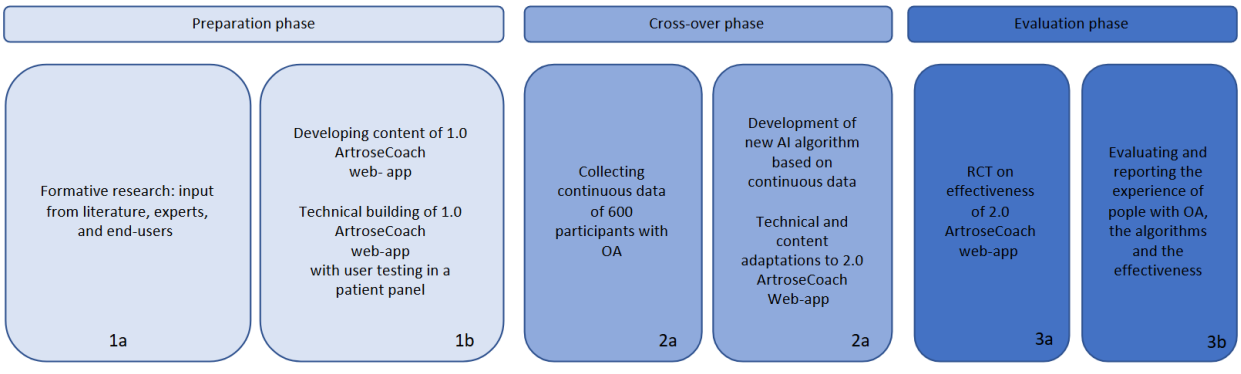

Figure 1: schematic overview of the research plan of SMART

Preparation phase: *In 1a information is collected from literature and experts to identify predictors for deterioration of pain and physical functioning in patients with OA. The content of the 1.0 ArtroseCoach web application is developed in co-creation with end-users and experts (1b) and with input from*

clinical guidelines for people with osteoarthritis. The 1.0 ArtroseCoach web application is tested for feasibility in a patient panel. Improvements are made in collaboration with end-users, refining the content of the app.

Observational cross-over phase: *In 2a (the current METC application), continuous data from 600 participants will be collected and used to develop a new (2.0) AI algorithm (the so-called data-driven models).*

Evaluation phase: *the effectiveness of the 2.0 ArtroseCoach app will be tested in a randomized controlled trial (3a) on self-management skills, physical functioning, participation, and use of care (stepped care compliance), and a process evaluation will be performed (3b). A separate METC application will be submitted for the evaluation phase of the SMART project.*

2. OBJECTIVES

Primary Objective

The primary objective of this cross-over study is to collect data to develop data-driven models to predict changes in physical functioning over time which will be used within the ArtroseCoach web application.

3. STUDY DESIGN

Study design

This study is a prospective longitudinal observational cross-over study of participants with hip and/or knee OA. Self-administered data and wearable data will be collected using the ArtroseCoach 1.0 web application. The collected data in this study will be used to develop a data-driven algorithm for the ArtroseCoach 2.0 app (Figure 1). Participants will be randomly assigned to one of the self-help programs within the ArtroseCoach web application (physical activity, weight management, sleep), or no intervention. The advice will be assigned at predetermined intervals (every 12 weeks), ensuring that no participant receives the same advice twice. The duration of the study for each participant from enrollment to the end of study is one year. Participants will be enrolled in 2025. Participants will be at their home setting while participating in this study.

4. STUDY POPULATION

Population

Participants with osteoarthritis on the hip and/ or knee will be recruited.

Inclusion criteria

In order to be eligible to participate in this study, a subject must meet the following criteria:

1. Have a hip or knee joint that, self-administered through a questionnaire, meets the National Institute for Health and Care Excellence clinical criteria for OA[13]:
    a. Aged 45 years or over and;
    b. Activity-related pain at the joint and;
    c. Joint morning stiffness that lasts no longer than 30 minutes or no morning stiffness at the joint;

2. History of pain at the joint for at least 3 months;
3. Have access to a smartphone with internet connection and an email address;
4. Able to give informed consent and willing to commit to all study evaluation and assessment procedures
5. Able to read and understand texts in Dutch at B1 level.

Exclusion criteria

A potential subject who meets any of the following criteria will be excluded from participation in this study:

1. Self-reported systemic arthritis (e.g., rheumatoid arthritis, gout);
2. Scheduled for lower limb joint surgery within the next year or underwent lower limb joint surgery (total hip, total knee) the last year;

Sample size calculation

We aim to include 600 participants in six months. During algorithm development, the data will be partitioned into multiple subsets to cross-validate the algorithms' performance. A random sample of the data will be used as test data to evaluate the algorithms' performance after they have been developed and optimized. This is necessary to test how well the models might perform on new data, for example, when the app is used by new users. We expect to better predict our outcome measurements (pain, physical functioning, and participation) by combining multiple predictors (such as age, BMI, and sleep quality) and by exploring non-linear relationships (where the effect of predictors like BMI may be stronger when changing in higher than lower values). Thus, it is crucial to have enough respondents to cover many combinations of characteristics. We do not yet know how many predictors will be important, the linearity of their effects, or the distribution of the predictors (e.g., how many people in our sample will report very poor sleep quality). However, we want to ensure that our sample size not only allows us to explore complex interactions and non-linear relationships but also to validate these relationships on unseen data. Typically, 20% to 30% of the data is used for testing the models on unseen data, with the remaining 70-80% used for training. In our case, this means using data from 100 to 150 respondents to test the models. We expect that this number will be sufficient to cover a range of multiple combinations of predictors (e.g., a young person whose BMI decreased from 30 to 28 and whose sleep quality improved from very low to low) multiple times within the test data. We are aware that more data will enable us to explore more complex relationships, which could result in better models. Our aim is to use the data to find an optimum between the amount of measurements required and model performance.

To account for potential dropout, we have calculated our sample size with a 15% attrition rate in mind, which means we plan to recruit approximately 600 participants to ensure that we have a final sample of approximately 510 participants after accounting for this expected dropout. Since the partners in this project have a large reach to potential participants and OA is the most common chronic disease, the likelihood of obtaining 600 participants is estimated to be high.

5. TREATMENT OF SUBJECTS

5.1 Investigational product

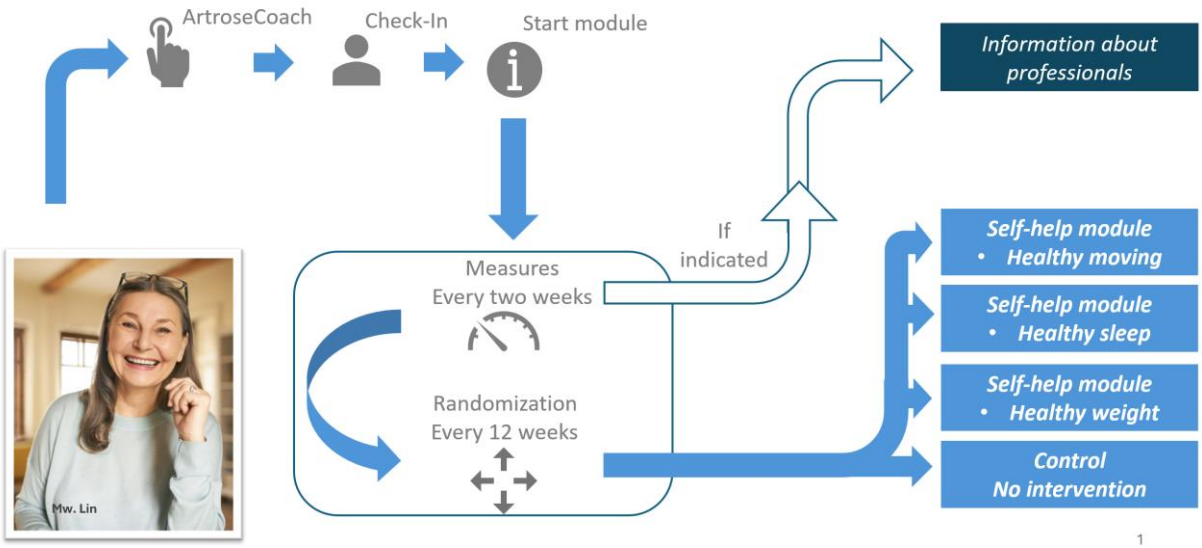

Figure 2: example of a patient journey through one year of using the ArtroseCoach web application.
.

ArtroseCoach web application

The ArtroseCoach 1.0 web application is primarily developed to collect data and utilizes expert-based algorithms to predict a worsening of levels of pain, physical functioning and participation. Besides collecting data, content in the ArtroseCoach is available and is complementary to OA clinical guidelines. Content includes self-help programs for physical activity, sleep or weight management, or information on healthcare providers that might be supportive at that moment. The allocation of these programs takes place through randomization among participants who meet the established safety criteria. An important safety requirement is excluding participants with extreme underweight (BMI <17.5) from being allocated to the weight management program. These individuals are at a high risk of malnutrition. In consultation with the Dutch Association of Dietitians, it has been deemed irresponsible to allow them to independently engage with a weight management program. Instead, they will be shown an informational message advising them to contact a dietitian. Participants who are underweight (BMI <18.5) and those with a PAR-Q score greater than zero will receive advice at the start of the cross-over study to consult a healthcare professional but will still be randomized to all treatment conditions of the study. People with extreme complaints (NPRS≥7) for four consecutive weeks receive information about relevant professionals. The ArtroseCoach 1.0 web application is accessible on participants' mobile phones and computers.

Based on the collected data in this cross-over study, data-driven algorithms will be developed for the ArtroseCoach 2.0 app and a future study. Patients are encouraged to regularly input relevant information into the app by filling in questionnaires about symptoms, level of physical functioning, psychosocial status, and any other data deemed necessary for the algorithms. The app generates

advice for each patient, taking into account their unique prognostic factors and lifestyle behaviours (sleep, physical activity and weight).
Note: this cross-over study is part of a large developmental project (Figure 1). Primary goal of this cross-over study is the collection of data to develop data-driven models, using a 1.0 version of the medical device. This medical device also provides advices, since the scored relevance of the advice will be measures in used in future machine learning models. In a next phase, the 2.0 version of the medical device will be evaluated in a randomized controlled trial study.

Content of the ArtroseCoach web application

The app features an intuitive and user-friendly interface to ensure easy navigation. Clear and concise instructions guide users through the input process, making it accessible to a wide range of individuals. All content is easy to understand at a B1 Level of Dutch. Different subgroups of patients (people with high and low digital health literacy) and HCPs tested the 1.0 prototype to ensure that the app technically works, is easy to use and relevant to patients with OA and HCPs. The content of the ArtroseCoach web application is based on national and international guidelines on hip or knee OA[3,14] . Existing educational materials and questionnaires are used and rewritten in B1 level. Patients with hip and/ or knee OA as well as B1 language experts provided feedback on the content.
The content of the app consists of different follow-up actions with education about the disease, lifestyle advice (e.g. exercise program, weight management program, sleep program) and healthcare professionals.

1. **Start program:** Every participant starts with the start program. The topics of the start program are information about: the disease, how pain occurs in OA, symptoms of OA, what you can do about your symptoms yourself, the importance of exercise and healthy lifestyle and general information about OA medication. The content is provided through text, videos and assignments.
2. **Physical activity program:** This program aims to increase the knowledge and level of physical activity (PA) and to improve muscle strength and stability of the hip and/or knee in twelve weeks. The movement program consists of three parts: information, a graded activity program (BGA) and strength exercises.
   Firstly, the information entails the influence of movement on pain, why 150 minutes of MVPA is important, chronic pain and handling energy. This information is provided to any participant included in the movement program.
   Secondly, the BGA program incorporates a baseline test, goal setting, time-contingent PA objectives (i.e., on fixed time points) and notifications to promote PA. An essential feature of the BGA program is the positive reinforcement of gradual PA, despite the presence of pain. The gradual increase in activities changes the perception that PA is related to pain and reinforces confidence to improve PA performance. The BGA intervention can be delivered with or without an activity tracker. In this study, 200 participants receive an activity tracker Fitbit Inspire) which they can connect tto the ArtroseCoach web application. The patient starts with a seven-day baseline measurement. After those seven days, the

following data is used: number of minutes of light physical activity (LPA) and moderate and vigorous-intensity physical activity (MVPA). The number of minutes of LPA and MVPA per week is seen as the baseline measurement. A personal goal is set for minutes of LPA and MVPA per week, aiming to surpass 150 minutes of MVPA per week after 12 weeks. The program will start at 75% of baseline measurement. Each week, the recommended amount of minutes MVPA is increased with ((personal goal – baseline) / number of weeks). Each day, the patient receives a positive reinforcement reminder of how far he/she is in reaching the week goal. Each week, the patient receives tailored feedback, based on the principles of graded activity.

Thirdly, exercises are provided through videos. The participant is encouraged to perform strength exercises twice a week for 20 minutes.

3. **Sleep program:** The sleep program aims to improve subjective sleep for people with clinical insomnia in 12 weeks. Insomnia is characterized by having trouble falling asleep, staying asleep and waking up too early. The sleep program consists of three parts.
   Firstly, there is a weekly educational theme about sleep or sleep hygiene. Sleep hygiene refers to a set of recommended behaviours a person can engage in throughout the day or before bedtime to promote good sleep. This includes abstinence from caffeine, alcohol, and nicotine late in the day, the practice of relaxation, regular exercise, regular sleep/wake times, modifying the environment (e.g., reducing impact of noise/light), no daytime napping, and minimal use of light-emitting devices (e.g., smartphones)[15].
   Secondly, participants are provided with mindful exercises (progressive muscle relaxation, mediation/ visualization). Mindfulness-based treatments are efficacious at reducing symptoms of insomnia and improving sleep quality[16].
   Thirdly, participants are encouraged to adjust sleep behaviours through setting goals, tailored feedback and prompts.
4. **Weight management program:** The weight management program is designed to help participants adopting a healthier diet and achieving a healthier weight if needed. Based on Body Mass Index (BMI) (kg/m^2), age, and ethical background it is determined whether participants have a healthy weight, under- or overweight, or extreme under- or overweight. For those with a healthy BMI it is mentioned that the weight management program can be used to receive advise for a healthy diet. Individuals who are under- or overweight are advised to work towards a healthier weight with support of the weight management program. While individuals who are extreme under- or overweight are recommended to seek professional assistance, but they are still able to utilize the weight management program.
   Participants are provided with a target weight, which is set at five percent below their current weight, because this amount can reduce disability in people with knee osteoarthritis[17]. They are then asked to complete the ‘Eetscore’, a short food frequency questionnaire evaluating diet quality across sixteen different food components[18]. After

completing the Eetscore, participants receive an overview with feedback for each food component, focusing on whether they consume an appropriate amount.
Throughout the program, six weeks are dedicated to providing tips for eating more or less from specific food components. Every week will focus on another food component. After participants received the feedback overview, participants are asked to select food components for which they want to receive tips. Most components can be addressed once, some can be addressed twice. The remaining six weeks focus on providing tips for weight loss. The twelve weeks alternate between focusing on food components and weight loss. Every day participants are provided a tip to aid them in their progress.
Additionally, on a weekly basis, participants receive informative texts covering different topics. Examples of topics include 'the relation between weight and pain from OA' and 'the significance of maintaining a healthy diet'.

5.2 Use of co-intervention

If participants start a treatment supervised by a healthcare professional, the participant can decide to stop or continue with the self-management program in the app. . Once a month, participants will be asked about the care they have received through a question.

5.3 Escape medication

Not applicable.

6. INVESTIGATIONAL PRODUCT

6.1 Name and description of investigational product(s)

General description: The ArtroseCoach web application translates general guidelines into feedback and suggestions regarding self-help programs on physical activity, weight-management and sleep, as well as relevant (healthcare) professionals. The application aims to support self-management for people with hip and/ or knee osteoarthritis.

Intended purpose: The ArtroseCoach web application supports individuals with hip and/ or knee osteoarthritis by monitoring their condition and identifying patterns that help to provide insights and just-in-time feedback. Coaching on physical activity, weight-management, sleep and giving advice to contact relevant (healthcare) professionals empowers people in their self-management. , The app aims to collect data to investigate if the advice was at the right moment. The application is used in the individual's daily life setting. Intended user: The intended users are individuals with hip and knee osteoarthritis

Sponsor = manufacturer: University Medical Center Utrecht, Postbus 85500
3508 GA Utrecht, 08875 555 55, info@umcutrecht.nl.

6.2 Summary of findings from non-clinical studies
Not applicable

6.3 Summary of findings from clinical studies
Within our research group, a cross-over study using ArtroseCoach will be performed. No previous clinical data is available for this device.

6.4 Summary of known and potential risks and benefits
The total benefit-risk analysis can be found in the Investigational Medical Device Dossier, chapter 5, Risk Management File ArtroseCoach.

6.5 Description and justification of route of administration and dosage
Not applicable

6.6 Dosages, dosage modifications and method of administration
Not applicable

6.7 Preparation and labelling of Investigational Medicinal Product
Not applicable

6.8 Drug accountability
Not applicable

7. NON-INVESTIGATIONAL PRODUCT
Not applicable.

8. METHODS

8.1 Study parameters/endpoints
Study parameters in this cross-over study can be categorized into main study parameters, secondary parameters and characteristics. Since this study aims to be inclusive with respect to individuals' level of digital health literacy skills, we have adapted all content (including questionnaires) of the ArtroseCoach in proficiency level B1. Level B1 ensures that approximately 80% of the Dutch population can comprehend the questionnaires thereby enabling participants with low digital health literacy to provide reliable data. Adaptation of the questionnaires into level B1 consisted of the following steps: 1) investigation of difficult words and long or passive sentences using the 'Klinkende taal' application and rewriting to less difficult words and shorter, active sentences; 2) rewriting of statements into questions, in accordance to the Pharos guideline for understandable questionnaires[19]; 3) feedback on understandability of questionnaires from our user committee.

8.1.1 Main study parameter/endpoint
The main outcome of the study is a dataset which allows us to train two machine learning models: one model that predicts deterioration of physical functioning at 12 weeks, and one model that gives advice for one of the three programmes (physical activity, weight-management or sleep program). To do so, a large number of knowledge-based factors need to be measured with a high frequency.

Physical functioning: will be evaluated using the Hip disability and Osteoarthritis Outcome Score (HOOS)[20](de Groot et al., 2007) or Knee injury and Osteoarthritis Outcome Score (KOOS)[21]. The Dutch HOOS has good reliability and construct validity for patients with hip OA [20]. It includes five subscales; pain, other symptoms, function in daily living, function in sport and recreation, and hip-related Quality of Life. The Dutch KOOS has good validity and reliability in patients with mild and moderate knee OA[21]. The HOOS and KOOS also evaluate joint stiffness.

8.1.2 [22]Secondary study parameters/endpoints

Pain: will be evaluated using the Numeric Pain Rating Scale.

Participation: will be evaluated using the Patient-Reported Outcomes Measurement Information System Experience[22].

Physical activity The BPAAT (or in Dutch: Vragenlijst Fysieke Activiteit (VFA)) consists of two questions regarding the frequency and duration of PA. By combining the results of both questions (scores can range from 0 to 8), the subject can be classified as insufficiently (0–3 score) or sufficiently active (>3 score)[23]. Insufficient refers to not meeting the WHO guideline of 150 minutes MVPA per week.

In a random subset (n = 200) of the study population, physical activity will be objectively measured using an activity tracker (Fitbit Inspire 3) worn in participants' home environments for the duration of the study. The primary aim of incorporating this device-measured active minutes is to evaluate whether the inclusion of objectively measured active minutes can increase the performance of the algorithms.

Weight will be evaluated using a scale in the participants' private environment. Self-monitoring of body weight ("self-weighing") has been emphasized as a key component of behavioural weight management programs. Two studies on overweight participants found that there was strong agreement between observed and self-reported weight values, suggesting self-reported weight values are a valid measuring instrument to evaluate weight[24,25].

Sleep quality is reported through the Pittsburgh Sleep Quality Index (PSQI). The PSQI is a widely used questionnaire to evaluate sleep and is recognized as valid and reliable[25].

Insomnia Severity Index (ISI) The ISI is a brief and valid instrument to detect insomnia and to evaluate treatment response[26]. A score between 8 to 14 indicates light sleep problems and a score of 15 or higher indicates clinically relevant sleeping problems. The minimal important difference was suggested at 6 points[27].

8.1.3 Other study parameters

Participant characteristics will be obtained (Table 1) with an intake questionnaire at baseline. Ethnicity (Asian/ non Asian) is asked because iBMI calculation differs for people with Asian ethnicity[28].

Relevance of each advice is evaluated on a Likert scale. Additionally, psychosocial parameters will be assessed prior to the movement, sleep or weight management program. The psychosocial parameters identified so far include attitude (e.g., "Do you find it important to [target behavior]?"), intention (e.g., "To what extent do you intend to [target behavior]?"), and perceived behavioral control (e.g., "I am confident that if I want to, I can [target behavior]").

Table 1

Measurement instruments

	Parameter	Measurement instrument
Characteristics	Date of birth	
	Sex (male/ female)	
	Weight (kg)	
	Length (centimeter)	
	Level of education	
	Smoking	
	Alcohol use	
	Comorbidity	
	Type osteoarthritis (hip/ knee)	
	Duration of osteoarthritis complaints	
	Ethnicity (Asian/ non-Asian)	
	Use of walking device	
	Use of pain medication	
	Health care utilization	
	Device for ArtroseCoach	
	Health skills	
	Digital skills	
Primary	Physical functioning	HOOS/ KOOS
Secondary	Pain	NPRS
	Participation	PROMIS-SF 8a
	Stiffness	HOOS/ KOOS
	Self-reported active minutes	BPAAT
	Device measured active minutes	Fitbit
	Fatigue	NRS
	Job satisfaction	NRS
	Depression and anxiety	HADS
	Experienced social support	NRS
	Kinesiophobia	BFOM
	Self-efficacy	ASES
	Insomnia severity	ISI
	Pain coping	PCI
	Weight	
	Physical activity readiness	PARQ

	Sleep quality	PSQI
	Psychosocial parameters to change movement, sleep or weight management behavior	
	Intention	
	Attitude	
	Perceived behavioral control	
	Engagement	Time spent with app opened per day

Note. Numeric Pain Rating Scale (NPRS), Hip disability and Osteoarthritis Outcome Score (HOOS), Knee Injury and Osteoarthritis Outcome Score (KOOS), Patient-Reported Outcomes Measurement Information System Short Form (PROMIS-SF), Brief Physical Activity Assessment Tool (BPAAT), Numeric Rating Scale (NRS), Hospital Anxiety and Depression Scale (HADS), Brief Fear of Movement (BFOM), Arthritis Self Efficacy Scale (ASES), Insomnia Severity Index (ISI), Pain Coping Inventory (PCI), Physical Activity Readiness Scale (PARQ), Pittsburgh Sleep Quality Index (PSQI).

8.2 Randomisation, blinding and treatment allocation

Participants who meet the safety criteria will be randomly assigned to one of the self-help programs within the ArtroseCoach web application (physical activity, weight management, sleep) or no intervention at predetermined time points (weeks 3, 15, 27, and 39). Each participant follows all four conditions but in a randomized order. There are 18 orders in which we are interested and thus people will be randomized to. Randomisation will be performed using a computer-generated allocation sequence to ensure an unbiased distribution of conditions across participants. Blinding is not applicable in this study, as participants actively engage with the assigned self-care module within the ArtroseCoach web application. Both participants and researchers are aware of the assigned intervention.

8.3 Study procedures

If a participant is included (the recruitment procedure is described in Chapter 11.2), the participant can download the ArtroseCoach app to their device or use the web-app. The app can be used on smartphones, tablets and computers. The causal contrast of this study concerns a comparison between no intervention for 12 weeks and the three different 12 week intervention programs. Every participant follows each of the conditions but in a random order. The four conditions can be ordered in 18 different ways. They begin their programs (or the control condition with no intervention) at weeks 3, 15, 27 and 39. Therefore, no one receives the same intervention twice. Extremely underweight participants (BMI < 17.5) do not follow the weight program due to safety concerns and instead have a period of 12 weeks of no program.

Baseline measurement

All outcomes as described in Chapter 8.1.2 are derived through the ArtroseCoach app.

Intervention

During the intervention, participants receive advice about their movement-, sleeping- or weight management behaviour through the ArtroseCoach app. The intervention is extensively described in Chapter 4.

Follow-up assessments

Follow-up assessments occur throughout the study. However, measurements are spread over all 52 weeks, resulting in 52 timepoints. Measurements are spread over a year to minimalize the load for participants, spreading these measurements does not influence the development of the AI model. Frequency of the measures are every two weeks for pain, physical activity and stiffness, for they are expected to differ faster than other outcomes. Other outcomes that are expected to change slower are measured every twelve weeks. Table 2 shows a schematic view of assessments that have a predetermined timepoint. Additional measures are done at the start and end of the movement-, sleep- and weight management program. Relevance of each advice and program is evaluated on a Likert scale. Pages can be favored using a thumb up/ thumb down sign. Finally, minutes of time reading per page will be automatically stored. All follow-up assessments are done remotely through the ArtroseCoach web application.

Table 2

Measurement timepoints

Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16-51***	**52**
General Measures																**…**	
Date of birth	**x**															**…**	
Sex	**x**															**…**	
Length	**x**															**…**	
Level of education	**x**															**…**	
Comorbidity	**x**															**…**	**x**
Type osteoarthritis	**x**															**…**	
Ethnicity (Asian/ non Asian)	**x**															**…**	
Use of walking device	**x**															**…**	**x**
Use of pain medication	**x**															**…**	**x**
Device for ArtroseCoach	**x**															**…**	
Duration of osteoarthritis complaints	**x**															**…**	
Health skills	**x**															**…**	
Digital skills	**x**															**…**	
Smoking	**x**															**…**	
Alcohol use	**x**															**…**	
Health care utilization	**x**				**x**				**x**				**x**			**…**	**x**
Sleep quality	**x**															**…**	

Primary															...	
Physical functioning			x											x	...	
Secondary																
Pain	x		x		x		x		x		x		x		...	
Stiffness	x		x		x		x		x		x		x		...	
Minutes physical activity	x		x		x		x		x		x		x		...	
Participation			x											x	...	
Job satisfaction					x									x	...	
Depression and anxiety					x									x	...	
Social support							x								...	
Kinesiophobia							x								...	
Pain coping							x								...	
PAR-Q									x						...	
Fatigue									x						...	
Self-efficacy									x						...	
Insomnia severity	x										x				...	
Weight	x										x				...	

*Pattern of week 1 until week 12 repeats from week 13 until week 51.

8.4 Withdrawal of individual subjects

Subjects can leave the study at any time for any reason if they wish to do so without any consequences by sending an e-mail to the research team. The investigator can decide to withdraw a subject from the study for urgent medical reasons.

8.4.1 Specific criteria for withdrawal

8.5 Replacement of individual subjects after withdrawal

Participants who withdraw from the intervention will not be replaced. They will be asked, without obligation, to complete the follow-up measurements to enable an intention to treat analyses. If a participant withdraws from the study for a reason not related to the intervention such as illness, a new participant will be recruited to reach the required sample size.

8.6 Follow-up of subjects withdrawn from treatment

Participants who withdraw for a reason related to the intervention will be asked, without obligation, to complete the follow-up measurements.

8.7 Premature termination of the study

No adverse events are expected. If multiple serious adverse events are recorded that influence the health of the participants, the study will be terminated prematurely.

9. SAFETY REPORTING

9.1 Temporary halt for reasons of subject safety

In accordance with section 10, subsection 4, of the WMO, the sponsor will suspend the study if there is sufficient ground that continuation of the study will jeopardise the subject health or safety. The sponsor will notify the accredited METC without undue delay of a temporary halt including the reason

for such an action. The study will be suspended pending a further positive decision by the accredited METC. The investigator will take care that all subjects are kept informed.

9.2 AEs, SAEs

9.2.1 Adverse events (AEs)

All AEs reported spontaneously by the subject or via answers to questions regarding side effects of the lifestyle advice in the app will be recorded. Besides, a button will be available at all time to report AEs. Participants who start the physical activity program answer the Physical Activity Readiness Scale and the advice are low risk, guideline-based advice, therefore the risk for AEs is estimated to be low.

9.2.2 Serious adverse events (SAEs)

A serious adverse event is any untoward medical occurrence or effect that

- results in death;
- is life-threatening (at the time of the event);
- requires hospitalisation or prolongation of existing inpatients' hospitalisation;
- results in persistent or significant disability or incapacity;
- is a congenital anomaly or birth defect; or
- any other important medical event that did not result in any of the outcomes listed above due to medical or surgical intervention but could have been based upon appropriate judgment by the investigator.

An elective hospital admission will not be considered a serious adverse event.

This cross-over study will include participants with hip and/or knee OA. We expect that our, low-risk, lifestyle intervention will not have any negative influence on the occurrence of SAEs. Participants will be asked to report any unexpected health complaints, including unexpected hospital admission or treatment in the app. When there are doubts about the link between the ArtroseCoach intervention and the occurrence of serious adverse events, these events will be reported as SAEs.

The sponsor will report the study related SAEs through the web portal *ToetsingOnline* to the accredited METC that approved the protocol, within 7 days of first knowledge for SAEs that result in death or are life-threatening followed by a period of a maximum of 8 days to complete the initial preliminary report. All other study related SAEs will be reported within a period of a maximum of 15 days after the sponsor has first knowledge of the serious adverse events.

9.2.3 Suspected unexpected serious adverse reactions (SUSARs)

Not applicable

9.3 Follow-up of adverse events

All AEs will be followed until they have abated, or until a stable situation has been reached. Depending on the event, follow-up may require additional tests or medical procedures as indicated, and/or referral to the general physician or a medical specialist.

SAEs need to be reported till the end of the study within the Netherlands, as defined in the protocol.

9.4 [Data Safety Monitoring Board (DSMB) / Safety Committee]

Due to the low risk of the study we deem a DSMB is not applicable

10. STATISTICAL ANALYSIS

Our primary research objective is to collect data to develop data-driven models which will be used within the ArtroseCoach web application to provide just-in-time personalized self-management advice. To this end, we will compare the performance of various machine learning models for each respective goal.

10.1 Machine Learning for Healthcare Data

Machine learning offers advantages over traditional statistical approaches by detecting nonlinear relationships and temporal dependencies in high-dimensional health-related data, which is crucial given that disease trajectories vary with time and prior outcomes. A major challenge in applying machine learning models to healthcare data is their lack of interpretability. A central aim of this study is to improve the transparency of machine learning models by incorporating expert knowledge and employing inherently transparent architectures[29] or using post-hoc explanation methods[30]. An additional advantage of incorporating expert knowledge is that it increases model robustness in the face of limited data[31].

Our analysis can be best summarized as two separate prediction tasks. We want to predict a) whether a patient's symptoms will worsen over the next period (personalized timing) and b) which intervention will provide the largest benefit to the patient (personalized advice).

10.2 Personalized timing

Users of our final Artrosecoach 2.0 app will receive advice for one of our three programmes (physical activity, weight-management and sleep). The timing of this advice will be based on whether the user is 1) already experiencing complaints above a certain threshold and 2) predicted to experience a negative development of their complaints in the upcoming period. To determine the right threshold for condition 1) we will consult the data that is collected during our cross-over study using Artrosecoach 1.0. Since participants start their programs with a wide range of complaints at baseline (at weeks 3, 15, 27 and 39), we can determine the treatment effects for patients with different pain levels at the beginning of the programs. This can help us to choose an appropriate threshold for program assignment for the final app.

To check whether condition 2) hold we will train a machine learning model to predict a worsening of symptoms at a 12-week follow-up. A worsening of symptoms is defined as a physical functioning (HOOS/KOOS) score of 62.5 or lower. To do so, the study design includes a 12-week period with no

intervention for each participant such that we can use this data to predict the development of complaints in the case of no intervention. If both conditions hold (i.e., the patient is above the threshold of complaints and our model predicts that an Artrosecoach 2.0 user will experience a physical functioning score below 62.5) they will be assigned to one of the three programs.
Performance on task a) of predicting whether patient will experience a worsening of symptoms benefits from taking the temporal dimension of the data into account. We aim to take temporal dependencies into account by employing recurrent neural networks such as Long Short-Term Memory or survival models. We can evaluate classification performance (will patient A experience pain symptoms above a score of X?) of Recurrent Neural Networks by using the calibration, Area Under the Curve or log-likelihood which offer added benefit over a simple accuracy metric by dealing with imbalanced data and accounting for the entire probability distribution. Survival models can be evaluated using the concordance-index which prioritizes patients that are in more urgent need for help. Which exact time-frame we will employ in the final model will be the subject of investigation by comparing the effectiveness of our models using different time-frames (i.e., predicting worsening of symptoms over the next two, four or eight weeks).
In addition, taking the temporal structure of the data into account allows for efficient use of the information our data collects, considering the limited sample size. The biweekly measurements over the course of one year will yield approximately 13,000 instances (500 patients*26 timepoints). This includes the expected dropout of 100 participants. In addition, expert knowledge, data augmentation, and regularization improve the robustness and interpretability of models under limited sample size.

10.3 Personalised Advice

For task b) we need to investigate which program will provide the largest benefit to the individual patient. A challenge in doing so is that we introduce confounding in the data because we are intervening on the patient's clinical outcomes through the use of our app and its respective programs. To tackle this confounding we want to employ a combination of randomisation and causal machine learning. A causal machine learning model will be trained to provide actionable personalized advice for users once the app determines a need for advice (personalized timing, see above section).
Causal inference requires the satisfaction of three conditions: exchangeability, positivity and consistency[32]. Randomization ensures that positivity is met due to participants having a nonzero possibility to be assigned to each of the conditions, except for extreme underweight patients for which a separate model that excludes the weight-management programme will be trained. Exchangeability (or unconfoundedness) requires that all confounders are measured. We based the selection of covariates on a literature review that aimed to identify all relevant predictors of pain, physical functioning and participation in patients with hip- and/or knee osteoarthritis[33]. Since empirical testing of the exchangeability condition is not possible, basing our covariate selection on previous literature provides confidence that we encompass the most relevant identified predictors (and thereby potential confounders) of OA.
The use of causal machine learning will allow us to detect heterogenous treatment effects (HTEs) and thereby offer personalized advice that is based on the user's patient characteristics. We do this by

analyzing counterfactuals, a common approach in causal inference[34]. Since we aim to provide actionable insights, we need to ensure that recommendations that imply causality in the form of "we expect treatment A to improve outcome B" are justified. This can be done by contrasting the actual event for a given patient (e.g., patient X took treatment A) with a prediction for the counterfactual case (patient X did not take any treatment). Methodologies such as causal survival forests can do this while incorporating right-censored data (i.e., data from patients for whom follow-up measures are missing)[35].

10.4 Use of sensory data

A subset of participants will wear a wearable so that we can use approximations of the minutes they spend performing MVPA. We chose the Fitbit inspire 3 fitness band for this purpose because this device connects to both Android and Apple smartphones, calculates a proxy measure for MVPA called *active zone minutes,* and provides a web-API, device API and software development kit for third party apps. We need to read data from the wearable at least daily, as we use this sensory data to provide treatment advice and predict the worsening of symptoms.

To automate such information exchange between the Artrosecoach app and the wearable in an automated way, we need to develop an Application Programming Interface (API) for Android and iOS devices respectively. Users create a personal Fitbit account and can then connect it to Artrosecoach 1.0 through a web API. This can only be done if users give their explicit permission. Users are free to delete their account and respective data from Fitbit whenever they decide to do so. Deleting data on Fitbit takes up to 90 days, after which the data cannot be recovered anymore. Whether they keep the wearable or not does not impact access to their data, as the data is retained on their accounts, not on their watch.

Data from the wearables is first stored on the respective users Fitbit account. Although Fitbit stores data in any of the countries they operate in, they comply with GDPR and legal mechanisms such as Standard Contractual Clauses (SCCs) to ensure that personal data transferred internationally is protected to EU standards. Additionally, in December 2020, the European Commission approved Google's acquisition of Fitbit with specific conditions. One key condition is that Google must not use health and wellness data from EU users for advertising purposes and must store such data separately from other Google data.

10.5 Reinforcement Learning

Reinforcement learning might be deployed for a small-risk task, such as optimizing the content and timing of reminder e-mails to increase the response rates of our participants. The model observes a respondent's current state (such as activity levels or behavioural change strategy) and chooses actions that aim to improve the user's response rate, for example see Hassouni et al[36]. The model then receives feedback in the form of new respondent data, which reflects the effectiveness of the chosen actions. This feedback loop helps the model to learn and adapt its strategies for better outcomes.

10.6 Evaluation

During algorithm development, we will utilize various machine learning techniques instead of testing hypotheses. As our focus is not on hypothesis testing, Type I error - typically associated with falsely rejecting a null hypothesis - is not a primary concern. Instead, our emphasis is on evaluating how well our algorithms perform on unseen data to test overfitting and ensure robustness.

We will use cross-validation or bootstrapping instead of a random split of the dataset for model development and internal validation. This will make use of the entire dataset, and by repeating all development steps within these techniques, a more reliable estimate of model performance and uncertainty will be made, as recommended in the TRIPOD statement[37]. This will allow for a more robust validation that better reflects performance in new samples from the same population. See section 10.2 for a specification of evaluation metrics for the given tasks.

10.7 Missing data

Participants are given 14 days to complete a questionnaire. If the questionnaire is not completed, they will receive a reminder email after 7 days. On the 14th day, a new questionnaire will appear. If it is determined that someone does not complete two consecutive questionnaires, they will be contacted by telephone. Despite these strategies, there will undoubtedly be missing data, especially considering the number of measurement moments.

Which data imputation method we will employ will largely depend on whether we assume the data to be Missing at Random (MAR) or Missing not at Random (MNAR). MAR means that the probability of a value being missing depends only on the observed data, while MNAR means that this probability depends on unobserved data (e.g., the value of the missing data itself). It is not possible to empirically distinguish between MAR (Missing at Random) and MNAR (Missing Not at Random), because MNAR relies on unobserved data. To rule out the effect of incorrect assumptions about missing data, we perform a sensitivity analysis and use MICE imputation to replace missing values where necessary. In addition, when participants decide to drop out of a program, we ask them to provide a reason from an extensive list of options. This further clarifies whether we capture all relevant sources of missing information.

Furthermore, loss of participants at follow-up can be accounted for by employing survival models, which classify such data as right-censored rather than missing. Survival models thus take the available data into account until the time of censoring (e.g., due to dropout) or the event of interest (i.e., the patient experienced a worsening or symptoms) occurred.

11. ETHICAL CONSIDERATIONS

11.1 Regulation statement

The study will be conducted according to the principles of the Declaration of Helsinki (version 13) and following the Medical Research Involving Human Subjects Act (WMO). Furthermore, the study will

adhere to the General Data Protection Regulation (GDPR) to ensure the protection of personal data and the Medical Device Regulation (MDR) where applicable.

11.2 Recruitment and consent

This cross-over study employs an open recruitment strategy aimed at reaching all individuals with osteoarthritis (OA) in the Netherlands. Recruitment efforts are primarily coordinated through the University Medical Centre Utrecht. To enhance outreach, specific strategies, such as the distribution of online and offline leaflets, will be utilized in collaboration with participating field labs, including Leidsche Rijn Julius Gezondheidscentra (LRJG) Utrecht, Stichting Gezondheidscentra Eindhoven (SGE), Fysiotherapeuten Vereniging het Gooi en Omstreken (FVGO), and Beweeghuis Maastricht. Additionally, advertisements in local newspapers, posters at bus stops, local coaches, lifestyle coaches, social media platforms, and through ReumaNederland, will be leveraged.

Given the fully digital nature of data collection and delivery in this study, obtaining remote digital informed consent is an appropriate approach. Individuals interested in participating can register via an open online Castor© form, where they provide consent to be contacted for a follow-up phone call. Alternatively, registration can be completed by contacting the research team via the phone number or email address provided on the study leaflet.
Upon completing the registration form, potential participants receive electronic informed consent documents. A follow-up phone call is scheduled at least one week after receipt of these documents. During this call, the investigator provides a detailed explanation of the study's objectives and procedures, addresses any questions, and verifies eligibility based on inclusion and exclusion criteria. To ensure comprehension, the teach-back method is employed, requiring participants to articulate their understanding of the study aims and procedures.
Eligible and willing participants digitally sign the informed consent form via Castor© e-Consent, a platform compliant with Dutch and international legal and regulatory requirements. Participants can download a copy of their signed consent form from the Castor© e-Consent portal. In cases where individuals experience difficulties with or reluctance toward the digital system, a paper version of the informed consent form is provided. Individuals may request a discussion with a member of the research team to address any concerns or questions at any moment.
If the individual is deemed ineligible or chooses not to participate, their personal information (e.g., name, contact details) will be promptly deleted. For those who do participate, the investigator will register them as study participants, and they will receive a link to a secure website (artroseapp.nl) for further registration.

11.3 Objection by minors or incapacitated subjects

Not applicable

11.4 Benefits and risks assessment, group relatedness

Participation in the e-COACH cross-over study is associated with a minimal risk. The burden for patients consists of the use of the ArtroseCoach application and filling out questionnaires at different time points during the year. The approximate time to complete the questionnaires at baseline is 30 min

and during the study 15 minutes. Participants with an indication to increase their physical activity, sleep quality or weight management will be advised to follow a program to improve these behaviours. The content of the programs is according to the OA guidelines. The exercises in the app are chosen in such a way that the risk of injury is minimized.

11.5 Compensation for injury

The sponsor/investigator has liability insurance which is in accordance with Article 7 of the WMO. Due to the negligible risks, an exemption from participant insurance has been requested.

11.6 Incentives

A total of 200 wearables will be randomly loaned to participants who have indicated their willingness to wear a smartwatch during the study. Participants can keep the smartwatch at the end of the study if they have completed 80% of the questionnaires and have regularly worn the wearables. Participants who do not receive a wearable will receive a gift voucher worth 15 euros at if they have completed 80% of questionnaires at half of the study and 15 euros if they have completed 80% of the questionnaires at the end of the study.

12. ADMINISTRATIVE ASPECTS, MONITORING AND PUBLICATION

12.1 Handling and storage of data and documents

Within the informed consent participants give permission to handle their data in accordance with the EU General Data Protection Regulation and the Dutch Act on Implementation of the General Data Protection Regulation. The handling of the data is documented in a secure online data management plan, which is checked for correctness by the data managers of the department at the University Medical Centre Utrecht. All original informed consents will be securely stored within the University Medical Centre Utrecht. The data collected by the web application will be stored in an electronic case record form (CRF, CASTOR). Data will be stored in a specific folder on the secure data environment of the University Medical Centre Utrecht to which only the primary researchers will access and will be treated confidentially. Each participant will be identified by a specific study number. The code is filed separately and will be available to the participating investigators for the duration of the study only. The CRF is available to the investigators and the subject involved and will not be disclosed to a third party. The subject will be informed in writing about these data storing and handling procedures, with a clear statement that discretion will be guaranteed. The administration of the study will be performed by the study investigators. Data are retrieved and stored according to GCP guidelines. The handling of personal data will be in line with the General Data Protection Regulation (in Dutch: Algemene Verordening Gegevensbescherming, AVG). *In accordance with the applicable regulations (Medical Device Regulations and WMO), all study-related data and documents will be retained for a minimum of 15 years after the completion of the study* The METC will be informed of the starting date of the study date of the study once the study commences.

12.2 Monitoring and Quality Assurance

An independent monitor (quality officer) will monitor the study data according to the regulations described under Good Clinical Practice (GCP). In part of the study subjects, the Informed Consents

are controlled. Additionally, during onsite monitoring, the officers will perform a Source Data verification of data described in the Case Report Forms to investigate the agreement between source data and study reports. The intensity of this verification is related to the study risk assessment. In particular, inclusion and exclusion criteria and the primary endpoints of the investigation are subject to monitoring. The monitor will evaluate whether (S)AE's are adequately reported within the time frame as directed by the Dutch law. Details of monitoring are described in the monitoring plan.

12.3 Amendments

Amendments are changes made to the research after a favourable opinion by the accredited METC has been given. All amendments will be notified to the METC that gave a favourable opinion.
All substantial amendments will be notified to the METC and the competent authority.
Non-substantial amendments will not be notified to the accredited METC and the competent authority but will be recorded and filed by the sponsor.

12.4 Annual progress report

The sponsor/investigator will submit a summary of the progress of the trial to the accredited METC once a year. Information will be provided on the date of inclusion of the first subject, the numbers of subjects included and number of subjects that have completed the trial, serious adverse events/ serious adverse reactions, other problems, and amendments.

12.5 Temporary halt and (prematurely) end of study report

The investigator/ sponsor will notify the accredited METC of the end of the study within 8 weeks. The end of the study is defined as the last participants' last measurement.
The sponsor will notify the METC immediately of a temporary halt of the study, including the reason of such an action.
In case the study is ended prematurely, the sponsor will notify the accredited METC within 15 days, including the reasons for the premature termination.
Within one year after the end of the study, the investigator/sponsor will submit a final study report with the results of the study, including any publications/abstracts of the study, to the accredited METC.

12.6 Public disclosure and publication policy

This study is financed by SIA (NWO), there will be no involvement of any kind in the public disclosure of the study's results. The results of this study will be disclosed unreservedly, according to the CCMO publication policy. Results of the study will be published. We expect to publish the results of the cross-over study around one year after the last measurements.

13. STRUCTURED RISK ANALYSIS

13.1 Potential issues of concern

For the total risk analysis, please see Chapter 5 of the IMDD, Benefit-risk analysis. For the detailed risk analysis, please see the Risk Management File.

Risks from all identified hazardous situations have been considered and evaluated. All residual risks are acceptable based on the criteria in the risk management plan. The residual risks are equivalent to the residual risks of similar medical devices available on the market.
Risk control measures have been reviewed. No new risks have been introduced by the risk control measures.

Benefits of this system:
During the ArtroseCoach intervention, the ArtroseCoach measures shows the lifestyle behaviors of the participant. This will provide the patients insight in and real-time feedback about the patients' behavior. The ArtroseCoach intervention is designed to improve lifestyle behaviors that are previously found to influence physical functioning and pain. All lifestyle advices are international guideline based. Using the ArtroseCoach could therefore lead to better patient outcomes.

Residual risks of this system:
The residual risks (all as low as possible) of the system are the following: *Failure of the system, Security breach, Incorrect measurement results, Misinterpretation of data, User cannot operate in ArtroseCoach, Misinterpretation of data by health care professionals.*

After implementing the specified mitigations, it is concluded that the remaining risks are acceptable when weighed against the benefits to the patient and are consistent with a high level of health and safety protection. These conclusions have been discussed with and accepted by Professor Dr. J.M.A. (Anne) Visser-Meilij in June 2024, the head of the department, who is not involved in this research project.

13.2 Synthesis
Risks from all identified hazardous situations have been considered and evaluated. All residual risks are acceptable based on the criteria in the risk management plan. The residual risks are equivalent to the residual risks of similar medical devices available on the market.
Risk control measures have been reviewed. No new risks have been introduced by the risk control measures.

14. REFERENCES
1.
Pharos. Begrijpelijke vragenlijsten – de basis voor goede zorg. 2022. (https://www.pharos.nl/infosheets/begrijpelijke-vragenlijsten-de-basis-voor-goede-zorg/).
2.
De Rijk MG, Slotegraaf AI, Brouwer-Brolsma E, Perenboom CWM, Feskens EJM, De Vries JHM. Development and evaluation of a diet quality screener to assess adherence to the Dutch food-based dietary guidelines. 2022;128(8):1615–25.
3.
LaValley MP, Lo GH, Price LL, Driban JB, Eaton CB, McAlindon TE. Development of a clinical prediction algorithm for knee osteoarthritis structural progression in a cohort study: Value of adding measurement of subchondral bone density. 2017;19(1).

4.
Caleyachetty R, Barber TM, Mohammed NI, Cappuccio FP, Hardy R, Mathur R, et al. Ethnicity-specific BMI cutoffs for obesity based on type 2 diabetes risk in England: a population-based cohort study. 2021;9(7):419–26.
5.
KNGF-richtlijn artrose heup knie. 2018.
6.
Bode RK, Hahn EA, Devellis R, Cella D. Measuring participation: The patient-reported outcomes measurement information system experience. 2010;91(9).
7.
Cijs B, Stekelenburg R, Veenhof C, Knoop J, Boymans T, de Rooij M, et al. Prognostic factors on changes in pain, physical functioning and participation in patients with hip- and/or knee OA : A systematic review. 2024;
8.
Lundberg SM, Allen PG. A Unified Approach to Interpreting Model Predictions.
9.
COPD K richtlijn. Vragenlijst fysieke activiteit [Internet]. 2008. (Meetinstrumenten in de zorg). Available from: https://meetinstrumentenzorg.nl/wp-content/uploads/instrumenten/Vragenlijst-fysieke-activiteit-meetinstr.pdf
10.
Li X, Li C, Zhang P. Predictive models of radiographic progression and pain progression in patients with knee osteoarthritis: data from the FNIH OA biomarkers consortium project. 2024;26(1):112.
11.
Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: A systematic review and meta-analysis. 2007. p. 433–9. (Annals of the Rheumatic Diseases; vol. 66).
12.
den Broeder L, Hilderink H, Polder J, Staatsen B, Dekker L. Publication Volksgezondheid Toekomst Verkenning 2024. Kiezen voor een gezonde toekomst [Internet]. 2024. Available from: https://search.datacite.org/works/10.21945/rivm-2024-0110
13.
Bossen D, Veenhof C, Karin,E.C. van B, Spreeuwenberg PMM, Dekker J, de Bakker DH. Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. 2013;15(11).
14.
de Groot IB, Reijman M, Terwee CB, Bierma-Zeinstra S, Favejee M, Roos EM, et al. Validation of the Dutch version of the Hip disability and Osteoarthritis Outcome Score. 2007;15(1):104–9.
15.
de Groot IB, Favejee MM, Reijman M, Verhaar JAN, Terwee CB. The dutch version of the knee injury and osteoarthritis outcome score: A validation study. 2008;6.

16.
Van Zaanen Y, Hoorntje A, Koenraadt KLM, Van Bodegom-Vos L, Kerkhoffs GMMJ, Waterval-Witjes S, et al. Non-surgical treatment before hip and knee arthroplasty remains underutilized with low satisfaction regarding performance of work, sports, and leisure activities. 2020;91(6):717–23.
17.
Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra S, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. 2019;27(11):1578–89.
18.
Ross KM, Eastman A, Wing RR. Accuracy of Self-Report Versus Objective Smart-Scale Weights During a 12-Week Weight Management Intervention. 2019;27(3):385–90.
19.
Yang M, Morin CM, Schaefer K, Wallenstein GV. Interpreting score differences in the Insomnia Severity Index: Using health-related outcomes to define the minimally important difference. 2009;25(10):2487–94.
20.
Friedman JH, Popescu BE. Predictive learning via rule ensembles. 2008 Nov 11;2(3).
21.
Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. 2015 Jan 7;350(jan07 4):g7594.
22.
Morin CM. The Insomnia Severity Index: Psychometric Indicators to Detect Insomnia Cases and Evaluate Treatment Response. 2011;34.
23.
Snijders GF, Den Broeder AA, Van Riel PLCM, Straten VHHP, De Man FHR, Van Den Hoogen FHJ, et al. Evidence-based tailored conservative treatment of knee and hip osteoarthritis: Between knowing and doing. 2011;40(3):225–31.
24.
Bica I, Alaa AM, Lambert C, van der Schaar M. From Real-World Patient Data to Individualized Treatment Effects Using Machine Learning: Current and Future Methods to Address Underlying Challenges. 2021 Jan;109(1):87–100.
25.
Ebner L. ExpertRuleFit: Complementing Rule Ensembles with Expert Knowledge [Software]. [Internet]. 2003. (Github). Available from: https://github.com/LuisaEbner/ERF
26.
el Hassouni A, Hoogendoorn M, Ciharova M, Kleiboer A, Amarti K, Muhonen V, et al. pH-RL: A personalization architecture to bring reinforcement learning to health practice. 2022.
27.

Safari R, Jackson J, Sheffield D. Digital self-management interventions for people with osteoarthritis: Systematic review with meta-analysis. JMIR Publications Inc; 2020. (Journal of Medical Internet Research; vol. 22).
28.
Wu C, Zou G, Chen M, Wan L, Kielmann K, McCormack B. Perceived challenges in delivering comprehensive care for patients following stroke: a qualitative study of stroke care providers in Guangdong Province, China. 2022;44(1):59–67. Available from: https://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=154689779&site=ehost-live
29.
Harvey-Berino J, West D, Krukowski R, Prewitt E, VanBiervliet A, Ashikaga T, et al. Internet delivered behavioral obesity treatment. 2010;51(2):123–8.
30.
Kloek CJJ, Bossen D, Spreeuwenberg PM, Dekker J, Dinny,H. DB, Veenhof C. Effectiveness of a Blended Physical Therapist Intervention in People With Hip Osteoarthritis, Knee Osteoarthritis, or Both: A Cluster-Randomized Controlled Trial Background. Integrating physical therapy sessions and an online application [Internet]. 2018. (Original Research; vol. 560). Available from: https://academic.oup.com/ptj
31.
Hiraishi M, Wan K, Tanioka K, Yadohisa H, Shimokawa T. Causal rule ensemble method for estimating heterogeneous treatment effect with consideration of main effects. 2023 Jul 28;
32.
Smink AJ, Van Den Ende CHM, Vliet Vlieland TPM, Swierstra BA, Kortland JH, Bijlsma JWJ, et al. "Beating osteoARThritis": Development of a stepped care strategy to optimize utilization and timing of non-surgical treatment modalities for patients with hip or knee osteoarthritis. 2011;30(12):1623–9.
33.
Irish LA, Kline CE, Gunn HE, Buysse DJ, Hall MH. The role of sleep hygiene in promoting public health: A review of empirical evidence. W.B. Saunders Ltd; 2015. p. 23–36. (Sleep Medicine Reviews; vol. 22).
34.
Osteoarthritis in over 16s: diagnosis and management NICE guideline. 2022.
35.
Rash JA, Kavanagh VAJ, Garland SN. A Meta-Analysis of Mindfulness-Based Therapies for Insomnia and Sleep Disturbance: Moving Towards Processes of Change. W.B. Saunders; 2019. p. 209–33. (Sleep Medicine Clinics; vol. 14).