Plan Overview

A Data Management Plan created using DMPonline

Title: e-cOAch cross-over

Creator: Femke Groen

Data Manager: Bas Vriens

Contributor: Pien Kempkes, Corelien Kloek, Mariette de Rooij

Affiliation: UMC Utrecht

Funder: Netherlands Organisation for Scientific Research (NWO)

Template: UMC Utrecht DMP with DPIA V.2

Project abstract:

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 cohort study 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.

Study design: This study is a prospective longitudinal observational study. Participants participate for one year with measurements every 2 weeks.

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 modules with education about the disease, lifestyle advice and behaviour change support (e.g., physical activity module, weight management module, sleep module).

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.

ID: 145769

Start date: 01-09-2023

End date: 29-08-2029

Last modified: 19-02-2026

Grant number / URL: KICH1.GZ03.21.005

Copyright information:

e-cOAch cross-over

1. General features

1.1. Acronym/short study title

e-cOAch cross-over

1.2 Division

- Hersenen (Neurosciences)

1.3 Department

Revalidatie, Fysiotherapiewetenschap & Sport

1.4 Path of the Research Folder

\\ds\GROUPS\HER\Onderzoek\Groep Pisters\24U-0004_e-cOAch_cohort

1.5 WMO/DEC

- WMO

1.6 METC number (only for human-related research)

Vidatum ethics code: 24U-0004

1.8 Research type(s)

- Use of questionnaires
- Clinical

1.9 Research design(s)

- Interventional

This study is a prospective longitudinal crossover study.

1.10 Mono or multicenter study (one choice)

- Monocenter

1.11 The role of UMC Utrecht is:

- Initiating / sponsor center

1.12 Which organization is the sponsor of the study?

UMCU

1.13 Name of datamanager consulted

Bas Vriens

1.14 Last check date by datamanager

2024-09-24

1.15 Indicate which laws and regulations are applicable for the project (please check all that apply)

- Wet Kwaliteit, klachten en geschillen zorg
- Gedragscode Gezondheidsonderzoek (Dutch)
- Medical Device Regulation
- Nederlandse gedragscode wetenschappelijke integriteit
- Wet Medisch-Wetenschappelijk onderzoek met Mensen (WMO) or Medical Research (Human Subject) Act
- Algemene Verordening Gegevensbescherming (AVG) or General Data Protection Regulation (GDPR)

2. Data Collection

2.1 Give a short description of the research data.

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format
Human	600	eCRF + cookies	nieuwe app*		.csv

*onderzoek heeft als resultaat nieuwe app. App wordt gebruikt om vragenlijsten uit te vragen en data wordt naar excel geëxporteerd. gebruikersdata wordt direct via app verzameld.

server: salut -> west europa

2.2 Describe the flow of the data (name systems used and/or third parties, recipients) <add link to location where diagram is stored in RFS>

"O:\Groep Pisters\24U-0004_e-cOAch-cohort\B_Documentation\7_QualityCheck\1\DMP.dataflow.SMART.pptx"

2.3 Estimated storage space for your project

- < 250 GB (e.g. questionnaires, textfiles, datasets)

2.4 Can you reuse existing data? If so, list the data source(s)

- No, please specify

We collect data through a new medical device. No data from Hix is used.

2.5 Describe how you will take care of good data quality.

Experimental data from patients will be collected through a web-application to an electronic Case Report Form (eCRF) in a certified Data Capture Tool: Castor. In the eCRF, skips and validation checks are built in. Data quality will be checked by an independent monitor from Julius Clinical. Data collection will be frozen before analysis.

#	Question	Yes	No	N/A
1.	Do you use a GCP-compliant Data Capture Tool or Electronic Lab Notebook?	X		
2.	Have you built in skips and validation checks?	X		
3.	Do you perform repeated measurements?	X		
4.	Are your devices calibrated?			X
5.	Are your data (partially) checked by others (4 eyes principle)?	X		
6.	Are your data fully up to date?	X		
7.	Do you lock your raw data (frozen dataset)	X		
8.	Do you keep a logging (audit trail) of all changes?	X		
9.	Do you have a policy for handling missing data?	X		
10.	Do you have a policy for handling outliers?	X		

2.6 Specify data management costs and how you plan to cover these costs.

#	Type of costs	**Division** ("overhead")	Department	Funder	**Other** (specify)
1.	Data Capture Tool license fee				RVB/UMCU
2.	Questionnaire license fee	X			
3.	Storage	X			
4.	Time of datamanager	X			
5.					

2.7 Please give some more details on other centers involved. What are the roles of the other centers involved? (what research activity does this organization carry out in relation to the study and the data?)

Organization	Role/research activity
UMCU	Sponsor
NWO	Funder + consortium member
Julius Clinical	Monitoring
Vrije Universiteit	Consortium member
Maastricht University	Consortium member
Hogeschool Utrecht	Consortium member

2.8 Which contracts are in place?

Organization	Contract Type with UMCU
Vrije Universiteit, Maastricht University, Hogeschool Utrecht, Salut (developer web application)	Consortium agreement; shared responsibility agreement

2.9 State how ownership of the data and intellectual property rights (IPR) to the data will be managed

UMC Utrecht is and remains the owner of all collected data for this study. The data is collected in a relatively large patient group and is very valuable for further, broader studies in Europe. Our data cannot be protected with IPR, but its value will be taken into account when making our data available to others, when setting up Research Collaborations and when drawing up Data Transfer Agreement(s).

Ownership of the data, publications and handling of personal data is described in the consortiumovereenkomst (SMART NWO-KIC), article 4.2 and 4.4.

2.10 Use of new technology. Does your study involve the implementation of a technology that has not been used before at UMC Utrecht?

- Yes

We will use a new medical device. This device is called ArtroseCoach web application and its purpose is used by individuals with a health-related complaint to monitor and modify their complaints through lifestyle coaching that aims to alter behaviour. These behaviours include movement, weight management, sleep behaviour. Changing these behaviours can lead to improved self-management. The application is used in the individual’s daily life setting.

2.11 Can this technology (incl. software) be seen as a medical device?

- Yes --> Contact MTKF for guidance

MTKF approved the use of ArtroseCoach for this study.

2.12 Will the study need/use personal data (directly or indirectly identifying)? For example, from the Electronic Patient Files (EPD; HiX), DNA, body material, images or any other form of personal data?"

- Yes. You have indicated that you are using personal data in your project. The following chapter is the Data Protection Impact Assessment (DPIA) for research data. It is derived from the full DPIA, in accordance with the privacy office of UMC Utrecht. Answering questions in this chapter helps to determine the risk of processing the personal data and what measures to take to minimize these risks.

3. Data Protection Impact Assessment (DPIA)

3.1 Are suppliers involved in the research project processing personal data from this study? (e.g. transcribe agencies, external laboratories, ICT helpdesk of eCRF, other EDCs (Castor, Redcap, Inform), DRE, Limesurvey, MS Forms)

- Yes

Castor is involved by processing data in this study

3.2 Is the supplier already contracted by UMCU?

- Yes

3.3 Are there any other centers or organizations involved in the research with which personal data are exchanged?

- No --> Skip to 3.6 if both 3.1 and 3.3 are No

3.4 Please indicate for each party involved in the dataprocessing, which role under the GDPR they have (controller, joint controller, processor, or sub processor)

Party involved	**GDPR role in relation to UMCU**	**Location (NL, within EEA (not NL), outside EEA)**	**Which security policy is in place?**
Hogeschool Utrecht	Joint controller	NL	ISO 27001
Vrije Universiteit	Joint controller	NL	ISO 27001
Maastricht University	Joint controller	NL	ISO 27001
Castor	Processor	EEA	ISO 27001 & NEN 7510
Anvil	Subprocessor	EEA	ISO 27001 & NEN 7510
Monitor	Processor	NL	ISO 27001 & NEN 7510

3.5 What type of sensitive personal data will be used?

- Racial or ethnic background
- Health data

Data about participants' sleep, diet and physical activity behaviour. Racial background (Asian, non-asian) will be used.

3.6 What type of directly or indirectly identifying personal data will be used? Indicate why you need this data. Is this truly necessary?

Type of personal data	Reason for collecting these data
Name	To contact participant in case of missing data
Address	Invullen
Telephone number	To contact participant in case of missing data
Email	To contact participant in case of missing data
Age (if fine grained)	Examine NICE criteria; Decribe research output; modeling of algorithms
Date (or month) of birth	Examine NICE criteria; Decribe research output; modeling of algorithms
Gender	Describe research output; modeling of algorithms.
Imaging e.g. MRI, pictures or video (can be health data)	NA
Sound recordings (may be health data)	NA
Location data (e.g. postal code)	Describe research output; modeling of algorithms
Personal interests	NA
Other -> describe below	
Weigth	Describe research output; modeling of algorithms.
Length	Describe research output; modeling of algorithms.
Level of Education	Describe research output; modeling of algorithms.
Health care utilization	Describe research output; modeling of algorithms.
Comorbidity	Describe research output; modeling of algorithms.
Use of walking device	Describe research output; modeling of algorithms.
Ethnicity (asian/ non-asian)	Determine BMI
Pain	Describe research output; modeling of algorithms.
Stiffness	Describe research output; modeling of algorithms.
Participation	Describe research output; modeling of algorithms.
Self-reported active minutes	Describe research output; modeling of algorithms.
Device measured active minutes	Describe research output; modeling of algorithms.
Sleep quality	Describe research output; modeling of algorithms.
Insomnia severity	Describe research output; modeling of algorithms.
Physical functioning	Describe research output; modeling of algorithms.
Depression and anciety	Describe research output; modeling of algorithms.
Self-efficacy	Describe research output; modeling of algorithms.
Pain catastrophizing	Describe research output; modeling of algorithms.
Fear avoidance beliefs	Describe research output; modeling of algorithms.
Central sensitization	Describe research output; modeling of algorithms.
Pain coping	Describe research output; modeling of algorithms.
Self-management skills	Describe research output; modeling of algorithms.
Stress	Describe research output; modeling of algorithms.

3.7 Select any vulnerable groups from which you will collect data?

- Patients

3.8 Which legally prescribed personal number will be used? Note: it is NOT allowed to use BSN (or its international counterpart) for scientific research purposes.

- None

3.9 Can the purpose of the study be achieved with anonymous or pseudonymized data?

- No, it is mandatory to be able to identify the data subject for my WMO study. When a study is subject to the WMO and subjects are included from participating sites, the directly identifiable data of the subjects, the informed consents and the key leading to the identity of the subjects are stored at the participating center. The Sponsor of the study receives only pseudonymized data. When the research requires direct identifiable data in the data, the dataset is stored in folder C_PersonalData of your research folder structure with access only for the persons that need access to this data.

3.10 Which measures are taken to prevent the data from being traceable to the natural person? Also consider the measures taken to prevent data breaches.

- We did a Risk Analysis for new software and applications used in this study.
- SOPs about who and how an employee has access
- SOPs about how to deal with a subject’s right on access, rectification, deletion and objection of their personal data
- Logging and monitoring on access to personal data
- Clear retention period(s)
- Parties have ISO27001 and/or NEN7510 certification(s)*
- 2FA/MFA before access to (health) data
- Encryption in case of data transfers
- Pseudonymization of data
- Role specific access to identifying data

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. An extensive report of the risk analysis is described in the Risk Management File.

3.11 Does the reuse of the data fit within the purpose for which they were originally

collected?

- Not applicable, we will not reuse data

3.12 What type of consent for using personal data is obtained?

- Study-specific or other type of Informed consent (e.g. broad consent, deferred consent).

3.13 What criteria, as formulated in the Dutch Medical Treatment Contracts Act (WGBO) and GDPR, is applicable for not obtaining informed consent?

- Not applicable, we ask consent or reuse data from an external provider.

3.14 Please explain why above mentioned ground for not obtaining informed consent is applicable for your specific study situation:

Not applicable, we ask consent or reuse data from an external provider.

3.15 Who will perform the objection-check and when?

Not applicable, we ask consent or reuse data from an external provider.

3.16 Check if all requirements, additional to the criteria in 3.13 as formulated in the Dutch Medical Treatment Contracts Act (WGBO), GDPR, and Gedragscode Gezondheidsonderzoek, are met:

- Not applicable, we ask consent or reuse data from an external provider.

3.17 Is there a dispute settlement or a party where the subject can go to with questions or complaints?

- Yes: this is described in PIF models and in objection explanation

3.18 Describe how you manage your data to comply to the rights of study participants.

- A subject can object to processing of their personal data or withdraw consent
- We inform the subjects about their rights of access, rectification and deletion of their data. In the information provision we describe the contact information in case a subject wants to exercise their rights,

3.19 Does the data collected concern data from which behavior, presence or performance (profiling) can be measured when this is not the purpose of the research?

- No

3.20 Are automated (i.e. without any human intervention) decisions made about the subjects based on the data?

- No

huidige project is info verzamelen tbv uiteindelijke ontwikkeling AI algoritme.

3.21 Describe the tools, procedures and transport methods that you use to ensure that only authorized people have access to personal data

- To send surveys, email address will be used in the EDC, but this is encrypted for the users in such a way that users can send emails to subjects without seeing the actual email address. No directly identifying personal data other than email address will be used in the EDC.
- We make use of a certified Electronic Data Capture (EDC) tool (Castor), with user roles defined in such a way that user accounts only have access to patients from own center with the necessary role to add, view, edit and export data, except for the sponsor of the study
- We use the secured Research Folder Structure that ensures that only authorized personnel has access to personal data, including the key table that links personal data to the pseudoID

3.22 Describe who will have access to which data during your study.

Data Access Table for Example answer 1

Type of data	Who has access
Direct identifying personal data	Project investigator team, Datamanager
Key table linking study specific IDs to Patient IDs	Project leader, Project investigator team, Datamanager
Pseudonymized data	Research team, Datamanager

4. Data Storage and Backup

4.1 Describe where you will store your data and documentation during the research.

The digital files will be stored in the secured Research Folder Structure of the UMC Utrecht. For analysis, data will be stored in DRE, because we need to collaborate with external colleagues. Paper dossiers will be stored safely in a locked cabinet in a locked room in the UMC Utrecht. A project specific procedure is in place for access to the paper dossiers. Documentation of this procedure is stored in the Research Folder Structure.

4.2 Describe your backup strategy or the automated backup strategy of your storage locations.

During data collection, backups will be made from the Anvil server and uploaded to the Research Folder Structure. Upon completion of data collection, all data are exported and saved in the Research Folder Structure where they are automatically backed up by the UMC Utrecht backup system.

5. Metadata and Documentation

5.1 Describe the metadata that you will collect and which standards you use.

We do not use metadata standards yet. The data from Anvil will be delivered including a data dictionary. For every variable this data dictionary contains an explanation of the values.

5.2 Describe your version control and file naming standards.

Files will be saved as a short description of the file and a date. Old versions will be saves in a folder "Oude versies"

6. Data Analysis

6 Describe how you will make the data analysis procedure insightful for peers.

I have written an analysis plan in which I state why I will use which data and which statistical analysis we plan to do in which software. The analysis plan is stored in the project folder, so it is findable for my peers. The analysis plan is also part of the protocol, which is planned to be published.

7. Data Preservation and Archiving

7.1 Describe which data and documents are needed to reproduce your findings.

The data package will contain: the raw data, the study protocol describing the methods and materials, the script to process the data, the scripts leading to tables and figures in the publication, a codebook with explanations on the variable names, and a ‘read_me.txt’ file with an overview of files included and their content and use.

7.2 Describe for how long the data and documents needed for reproducibility will be available.

Data and documentation needed to reproduce findings from this WMO study with a medical device will be stored for at least 10 years.

7.3 Describe which archive or repository (include the link!) you will use for long-term archiving of your data and whether the repository is certified.

After finishing the project, the data package will be stored at the UMC Utrecht Research Folder Structure and is under the responsibility of the Principal Investigator of the research group. The (meta)data will be published in DataverseNL, the preferred UMCU repository.

7.4 Give the Persistent Identifier (PID) that you will use as a permanent link to your published dataset.

I will be using a DOI-code and will update this plan as soon as I have the code.

8. Data Sharing Statement

8.1 Describe what reuse of your research data you intend or foresee, and what audience will be interested in your data.

The raw data can be of interest for other researchers or for spin off projects.

8.2 Are there any reasons to make part of the data NOT publicly available or to restrict access to the data once made publicly available?

- Yes (please specify)

As the data is privacy-sensitive, we publish the descriptive metadata in the data repository with a description of how a data request can be made (by sending an email to the corresponding author). In the event that peers like to reuse our data this can only be granted if the research question is in line with the original informed consent signed by the study participants. Every application therefore will be screened upon this requirement. If granted, a data usage agreement is signed by the receiving party.

8.3 Describe which metadata will be available with the data and what methods or software tools are needed to reuse the data.

All data and documents in the data package mentioned in 7.1 will be shared under restrictions.

8.4 Describe when and for how long the (meta)data will be available for reuse

- (Meta)data will be available as soon as article is published

8.5 Describe where you will make your data findable and available to others.

We will use DataverseNL as a repository for our research data, we will follow the UMC Utrecht guidelines for publishing research data.