DMP title

Project Name AVI-DYS (FWO DMP) - DMP title

Project Identifier 3M210649

Grant Title MSCA SoE fellowship_12ZZW22N

Principal Investigator / Researcher Helga Haberfehlner

Description AVI-DYS consists of two parts: a retrospective part (AVI-DYS retro) making use of available videos and a prospective part (AVI-DYS home) exploring the feasibility of using the methodology in less standardized situations (i.e. home) Abstract: Movement disorders in dyskinetic cerebral palsy (DCP) are associated with impaired muscle tone regulation and interfere with intentional movements. To treat movement disorders invasive neuromodulation treatments are increasingly applied within DCP. Effective monitoring is extremely important for the indication, evaluation and dosing of these interventions. Current methods to assess movement disorders in DCP are insufficient and time consuming. In addition hospital measurements are not representative of the real-word situation, as movement disorders may vary considerably during the day, and increase with emotions or pain. Therefore, the aim of the study is to develop an objective measurement technique that can be used in everyday situations. We suggest using supervised machine learning to automatically classifies dyskinetic movement patterns and assess severity, using data extracted by markerless motion from videos. Markerless motion tracking runs already automatically for common movements of abled bodied persons. We will 1) re-train an existing human model on about 3000 unique video sequences of 130 children with DCP performing everyday tasks and 2) train algorithms that maps features from the videos to clinical scoring. The data consists of the raw videos, the extracted stickfigure data from markerless motion tracking and the clinical scoring.

Institution KU Leuven

1. General Information

Name applicant

Helga Haberfehlner

FWO Project Number & Title

MSCA – SoE fellowship_12ZZW22N

Automated VIdeo-based assessment of DYSkinesia in cerebral palsy using markerless pose estimation and machine learning (AVI-DYS)

Affiliation

- KU Leuven
- Other

Amsterdam UMC

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data
- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Data for **AVI-DYS retro** will use **existing data of 130 children and young adults** with dyskinetic cerebral palsy (CP) (age 4-24 years). The data consists of video sequences and appertaining clinical scores of dystonia and choreoathetosis scored on the Dyskinesia Impairment Scale (DIS). The data has been collected at four different centers (KU Leuven, Amsterdam UMC, Maastricht UMC+ and University Hospital Cologne) for the following purposes: (1) to develop and evaluate the DIS and instrumented measurments of dyskinesia (data KU Leuven), (2) to assess the effect of intrathecal baclofen (IDYS trial) (data Amsterdam UMC and data Maastricht UMC+) and (3) to assess the effects of deep brain stimulation (data University Hospital Cologne). The available videos are all recorded following the standardized DIS video protocol including

sequences of 30 seconds in different postures (lying, sitting and standing), different views (frontal and profile view, general view and close-up), during different activities (e.g. rolling, grasping and speaking) and during rest (see Tabel 1 for an overview of the data).

Table 1: Estimation of available Dyskinesia Impairment scale (DIS) data:

Research Site	Original study	Participants (DIS data* x averaged sequences to be used**)	Video sequences (Range of minutes that can be processed
KU Leuven	DIS development , Move-IT, IDCA	83 (113 x 12)	1.356 (452-678)
Amsterdam UMC	IDYS-trial	34 (94 x 15)	1.410 (470-705)
Maastricht UMC+	IDYS-trial	7 (13 x 15)	195 (65-98)
University Hospital of Cologne	STIM-CP	4 (16 x 10)	160 (53-80)
	Total	**128 (236 x 12)**	**3.121 (1.022-1.534)**

* some participants are measured severeal times; ** not all sequences can be used / or are measured, Dyskinesia Impairment Scale (DIS), Instrumented dystonia and choreoathetosis assessment protocol (IDCA), Intrathecal baclofen treatment in dystonic cerebral palsy (IDYS), bilateral pallidal stimulation in dyskinetic cerebral palsy (STIM-CP)

Next to the DIS score the videos will be scorde on the Barry-Albright Dystonia scale (BADS) and the videos will be split in timewindows of five seconds and re-scored. From the videos x,y-coordinates of the following bodypoints will be extracted by markerless motion tracking: forehead, chin, shoulders, elbows, wrists, midhand, hips, knee and ankles, midfoot. The following patientcharacteristics will registered: Type CP (dyskinetic, dyskinetic-spastic), age, sex, Level of Gross Motor Classification System (GMFCS), Level of Manual Ability Classification System (MACS), treatment (if pre-post videos), time point of video. In Tabel 2 an overview of the data used for AVI-DYS retro including format and volume estimation.

Table 2: Overview data AVI-DYS retro

Type of data	Format	Volume estimation	How created
Raw video data	.avi	about 3100 sequences x 15.000 kB/sequence = about 50 GB	existing; overview see table 1
X,y-coordinates	.csv	from about 3100 sequences x 580 kB/sequence = about 2 GB	extracted by DeepLabCut
DIS-clinical scores	.csv	~10 MB	available, as database in spss or excel from each study side
DIS-clinical scores in 5 seconds time-windows	.csv	~10 MB	scored by clinical rater by watching videos
BADS-clinical scores	.csv	~10 MB	scored by clinical rater bij watching videos
Patientcharacteristics: Type CP, age, sex, GMFCS, MACS, treatment, time point of video	.csv	<1MB	Data extraction from previous trial registration, registered in RedCap

For **AVI-DYS home new data of 20 children and young adults** with dyskinetic CP aged between 4-24 years will be measured. Two protocols will be performed the standardized DIS video protocol and a natural environment protocol, mimicking home-based measurements. Overview of the data is provide in Table 3.

Table 3: Overview data AVI-DYS home

Type of data	Format	Volume estimation	How created
Raw video data DIS	.avi	about 15 sequencesx 20 subjects x 15.000 kB/sequence=about 5 GB	videorecorded during trial
Raw video data natural protocol	.avi	about 20 minutes x 20 subjects=about 20 GB	videorecorded during trial
X,y-coordinates DIS	.csv	from 15 sequences x 20 subjects x 15.000 kB/sequence=>200 MB	extracted by DeepLabCut
X,y-coordinates natural protocol	.csv	about 20 minutes x 20 subjects = >1GB	extracted by DeepLabCut
DIS-clinical scores	.csv	~10 MB	scored by clinical rater by watching videos
DIS-clinical scores in 5 seconds time-windows	.csv	~10 MB	scored by clinical rater by watching videos
BADS-clinical scores	.csv	~10 MB	scored by clinical rater bij watching videos
natural protocl scored in 5 seconds time-windows	.csv	~10 MB	scored by clinical rater bij watching videos
Patientcharacteristics: Type CP, age, sex, GMFCS, MACS, treatment, time point of video	.csv	<1MB	Data extraction from medical record registered in RedCap

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- Yes

Privacy Registry Reference:

G-2022-4767 (for AVI-DYS home)

for AVI-DYS retro: PRET need to be performed

Short description of the kind of personal data that will be used:

Personal details (e.g. age, gender)
Video recordings
Health: diagnoses and symptoms (Type CP, GMFCS, MACS, clinical scoring dystonia and choreoathetosis, treatment)

Personal information for contact purposes (e.g. name, address, phone number, e-mail), which will not be used in any further analysis. Participants will be asked whether this information can be stored in a database for sending the newsbrief and for future research, via a separate informed

consent procedure in accordance with the General Data Protection Regulation.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- Yes

AVI-DYS retro: use of data from humans
AVI-DYS home: experiments with humans (observational)
For both ethical approval is required, not approved yet.
AVI-DYS home: S66357 (submitted EC)
AVI-DYS retro:
Ethical approval for study sites in Germany and the Netherlands will be locally arranged.

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

Projectdocumentation (without data) will be on project folder on Oneddrive (shared with project team) with a document explaining the folder structure. Standard operating procedure (SOP) will be stored as pfd in the project folder, including research methods, practices, instructions given to participants, etc., as well as a blank copy of the information letter and informed consent form. The SOP will also include the link to the datastorage on the secure KU Leuven networkdrive and the location of the data at the Amsterdam UMC.

The code will be available on Github, with a detailed readme. Syntax of SPSS of statistical analysis will be stored.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

Metadata will be created to make the data easier to find. For the majority of that data, metadata will be provided as readme, word or excel files, containing all settings and technical descriptions of the experiment and data. Data documentation is saved in the same folder in which the corresponding datasets are saved, for example as a Readme.txt file.

5. Data storage and backup during the FWO project

Where will the data be stored?

REDCap will be used to capture study related data. The following data will be collected in the eCRF: Type CP and location, age, sex, GMFCS, MACS, DIS and BADS scoring, scoring within 5 seconds time window.

Video data (privative sensitive) from the KU Leuven and the University Hospital of Colognes will be stored on the secured KU Leuven's secure internal servers and processed locally. Only the investigators will have access to these files. The PI of this project (Elegast Monbaliu) will be the only one who can grant access to this network drive.The video data of the Amsterdam UMC and Maastricht UMC+ will remain within the Amsterdam UMC (storage for IDYS-trial) and processed locally. The PI of the Amsterdam UMC (Annemieke Buizer) is responsible for the storage, backup and archiving of these data and access to the files.

Processsed data of all studysites: scoring of the videos and extracted x,y-coordinates as will be uploaded to RedCap as .csv file

How is backup of the data provided?

For RedCap required back-up procedures are in place.

Back-up of the data on KU Leuven secure network drive will be provide with automatic daily backup procedures. Additionally, a mirror of the data is provided in a second ICTS data centre for business continuity or disaster recovery purposes.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

Sufficient storage and backup capacity for the data as described in part 2 of this DMP is provided on the KU Leuven servers and networks, and participatin sites. This is about 50GB for the KU Leuven and about 30 GB for the Amsterdam UMC.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

AVI-DYS data will be stored on the L:/drive Revaki Campus Brugge in an project folder (AVI-DYS). About 50 GB of data storage is needed. On the L:/drive 5 Tb are available. The costs of €641,95/year/5 Tb for the storage on the L:/drive are covered by the reserach group.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Participant identities will be linked to data using an alphanumeric code. The separate and uniquely double pass-word coded "Subject Identification Code List", which matches identifying codes with the subjects' names, will be managed by the principle investigator (EM) and stored separately, using the Digital vault for private data service of the ICTS, KU Leuven. This system involves a secure and operating system in ICTS's special, secure environment for private data.

The participant code of the other study sites will be kept at the local sites.

Access via RedCap will only be provide to authorized persons of the study team.

Acess to the video data on the secured KU Leuven's secure internal servers can only be granted by the PI of this project (Elegast Monbaliu). Only the investigators of the studygroup will have access to these files.

Access to the video data of Amsterdam UMC is granted by the local PI of this project (Annemieke Buizer)

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

Final processed data from the all studiesites as well as the raw data of the KU Leuven will be stored for minimal 10 years.

Archiving for the raw data on the Amsterdam UMC server is the responsibility of the local PI, concerning regulation data will be kept at least 15 years.

Where will the data be archived (= stored for the longer term)?

The data will be archived on the university's central servers (with automatic back-up procedures) (K-drive). The project documentation will also archived on the K-drive (i.e. transferred from the one-drive as used during the project).

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

For the final year of the project an additional part in the budget is foreseen for data archiving (K-drive) for a period of minimal 10 years. The expected costs are €12,84/year/100Gb. Costs related to archiving will be covered for 50% by the Department of Biomedical Sciences and for 50% by the project.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- Yes. Specify:

Raw data (i.e. videos) cannot be shared public due to privacy restrictions.

Which data will be made available after the end of the project?
The extracted x,y-coordinates, clincial scoring and selected patientcharcteristics will be published in a public available Research Data Repository (RDR) of the KU Leuven) under a CC-BY license.

Where/how will the data be made available for reuse?

- In an Open Access repository
- Other (specify):

Anonymised data will be published in a csv format in the RDR of the KU Leuven together with documentation.
The code for processing the data will be released on Github and made citable with zenodo repository.
The trained model for tracking of bodyparts and the models for prediction of dystonia and choreoathetosis will be published on zenodo.
Raw data video will be eventually shared within the European Network of Dyskinetic Cerebral Palsy upon request (if parents/participants have consented).

When will the data be made available?

- Upon publication of the research results

Who will be able to access the data and under what conditions?
During the project:

- All involved researchers

After the project:
The data published in RDR will be accessable under a CC-BY license.

What are the expected costs for data sharing? How will the costs be covered?
Data will not exceed 50 GB, storage up to 50 GB is possible in th RDR of the KU Leuven.

8. Responsibilities

Who will be responsible for data documentation & metadata?
The postdoc researcher (Helga Haberfehlner) associated with this project will be responsible for data documentation & metadata, under supervision of the PI.

Who will be responsible for data storage & back up during the project?
The postdoc researcher (Helga Haberfehlner) associated with this project will be responsible for data storage & back up, under supervision of the PI.

Who will be responsible for ensuring data preservation and reuse ?
The PI (Elegast Monbaliu) will be responsible for ensuring data preservation and reuse.

Who bears the end responsibility for updating & implementing this DMP?
The PI bears the end responsibility of updating & implementing this DMP.