Unravelling the early OA cartilage 'mechanome' as basis of regenerative mechano-therapy - an in vitro, in silico and in vivo approach

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: ***N****(ew data) or* ***E****(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
1	Informed Consent form	N	D	T	.txt, .pdf	<1GB	
2	Finite Element simulations	N	D	M, SO	Numerical software specific data (.odb, .inp)	<1GB	
3	MRI	N	D	I	Numerical software specific data (Dicom format .DCM)	<1GB	
4	Segmented cartilage meshes	N	D	M	Numerical software specific data (.stl)	<1GB	
5	Cartilage sampels	N	P	Cartilage tissue, human cells			Harvested from human cartilage tissue explants
6	Histological and cell culture analysis	N	D	I	(.tiff, mp3, .pdf)	200-300 MB	
7	Gel electrophoresis, Westrn Blot	N	D	I	(.tiff, mp3, .pdf)	1 GB	
8	RNA sequence analysis	N	D	Molecular assessment	(.bam, FASTQ,.xls)	<100 GB	
9	Bioreactor loading data	N	D	Numerical sofware specific data	(.csv)	500 MB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

NA

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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- Yes, human subject data (Provide SMEC or EC approval number below)

- Yes, animal data (Provide ECD reference number below)

Cartilage explants will be covered by an amendment to *S61930*
Animal experiments will be covered by amendment to *P134/2018*

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- Yes (Provide PRET G-number or EC S-number below)

S61930
For Data source 5, waste material from endoprosthesis will be collected. Surgeons will provide data on age, gender, and reason for surgery. All data will be pseudonymised.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- Yes

Data from this project may considered to claim intellectual property rights on the advise of Leuven R&D's valorisation team. LRD will be responsible for patent management and eventual licensing. Data may be used for industrial colla

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

The following documentation will be provided:
(1) a table of content (excel file and csv) with all project-related experiments including experiment number, date of implementation and name of the researcher who stored the experiment,
(2) a brief description of the goal of the experiment and related work package (word and txt file),
(3) a detailed protocol or link to an existing standard protocol (SOP) which will enable other researcher to repeat the experiment,
(4) all data or link to another file with the (raw) data,
(5) samples that are generated during the experiments and will be stored and listed in an excel file,
(6) if appropriate, illustrations of the data with legends and statistical analysis. In case documentation is written or available in notebooks or stored in other files, a link will be provided.

With the help of these documentations, every authorized researcher will be able to look up all the information about the experiments performed and repeat the experiment in precisely the same way.
All data will be coded. This will consist of:

- Approved Ethical Commission: description of study protocol (.pdf)
- Experimental protocols: description how the data are collected and generated (software, materials, set-up, settings (.docx) and how data are processed (software, protocol, guidelines, …) (.docx, read.me text files). For the biological experiments, detailed protocols or link to existing standard protocols (SOPs) will be provided, which will enable other researcher to repeat the experiments.
- Measurement forms: notes during data collection (printed paper)
- Raw experimental data: storage of original physical data and folders with original digital data in software-specific files
- Processed data: folder with digital data in the software-specific files, spreadsheets with results (.CSV, .xls)
- Patient identifier record: subject study code (.xls) This patient record file is the only document that provide the link between the study code of the patient and patient's personal data.
- Subject recruitment files: only subject study code, personal data (for example, age, gender) reason for surgery. The subject recruitment files described the measurements info for each patient, whereby the patient's identiy is coded.
- The patient identifier record (PIR) will be stored separately on another location than the subject recruitment files, using the service from UZ Leuven (password protected network location) to keep private data safe and this is supervised by the PI.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

For most of that data, metadata will be provided as readme, word or excel files, containing all settings and technical descriptions of the experiment and data.
More specifically, Our anticipated data collection does not involve setups for which a metadata standard and specific repository exists. Data will, therefore maximally be stored in portable formats such sc txt, csv and pdf
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.
Raw experimental data (from the bioreactor) are managed on a software-specific data management platform.
Readme files and logbooks will be generated to describe the different steps taken in the simulation workflow.
For imaging data (i.e. MRI and microscopy), a large part of the metadata is included in the header files of the original images. These files contain information regarding the acquisition settings (e.g. acquisition time, flip angle, TE, TR, field of view, slice thickness).

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Personal network drive (I-drive)
- Large Volume Storage

Raw and processed physical and digital data will be collected per assessment.
Physical data derived from cartilage samples will be stored in freezers at the Skeletal Biology and Engineering Laboratory. Paper lab notebooks on biological experiments will be stored here as well.
Digital data will be stored in a large volume storage (J-drive) of the KU Leuven, specifically developed to store large amounts of data for long periods.
Additionally, copies can be made on the individual work pc of the researchers involved in the project.
MRI images will additionally be stored and backed up on the dedicated servers of MOSAIC.
The paper copies of the descriptive data will be stored in a secured locker at the Department of Rehabilitation Sciences, Building The Nayer, of the KU Leuven. Only authorized personnel will have access to this locked storage room as they must be granted access by the PI (Ilse Jonkers).

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

The paper copies will be digitized and together with the digital data stored on the university's secure network drive with automatic daily back-up procedures. Additionally, a mirror of the data is provided in a second ICTS data center for business continuity or disaster recovery purposes.
Cartilage tissue samples and the analysed samples will be kept for 5 years postproject in freezer (temperature ranging from -20° till -150°, depending on the sample type), at the Laboratory of Skeletal Biology and Engineering.
Digitial data automatically stored on the acquisition laptop during data collection, will be manually transferred via external hard drive to the secure servers. This external hard drive is provided as automatic back-up of the acquisition laptop.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

Sufficient storage and backup capacity are available at KU Leuven

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

All data will be stored in a protected environment. Research data can only be accessed by a login following KU Leuven's policy for identifier and with pasword. The digital, pseudonymised, data (i.e. coded and containing no personal information) will be stored in a secure university environment. The PI of this project (Ilse Jonkers) will be the only one who can grant access to this network drive. The separate and uniquely double pass-word coded "Subject Identification Code List", which matches identifying codes with the subjects' personal details, will be managed by the principle investigator (Ilse Jonkers) 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."

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

Terabyte storage is anticipated as a need and will be covered by the grant (approx 520€/terabyte/year).

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- KU Leuven RDR
- Shared network drive (J-drive)
- Large Volume Storage (longterm for large volumes)

Digital data will be archived on the secured university's network drive. Additionally, data will be stored offline on two external hard drives when the project is finished. Hard copies (eg. the Informed Consent forms, measurement forms and paper lab notebooks) are kept in locked cabinets in the lab of the PIs concerned.
The physical cartilage samples will be stored for a long term in freezer (temperature ranging from -20° till -150°, depending on the sample type), at the Laboratory for Skeletal Biology and Engineering

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

For this project data storage of 1Tb is anticipated, resulting in a cost of 520 euro per year, that can be covered by the grant and beyond.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- Yes, as open data
- Yes, as restricted data (upon approval, or institutional access only)

The full anonymized dataset will be made available after publication of the data and upon request with the PI. However, valorization initiatives may prevent the open sharing of the data.

If access is restricted, please specify who will be able to access the data and under what conditions.

If applicable, data sharing agreements will be elaborated wit KU Leuven Research & Development

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

Please explain per dataset or data type where appropriate.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other (specify below)
- KU Leuven RDR (Research Data Repository)

During the project as well as after the end of the project, the published data will be available via an open access repository (e.g RDR) and upon request by email to the PI. These published data contain the results of processed coded data presented in tables. Reference databases for medical imaging and microscop will be established by the end or after the end of the project. As part of the valorisation plan, these databases maybe put available for external users through open source pathways. In that case, these data will be made available after appropriate IP protection. Importantly, only data of participants who granted their approval for re-use, either within the research group (closed data) or outside the research group (open data), will be made available (also see 'Who will be able to access the data and under what conditions?') Patient-specifc data will only be shared ensuring the privacy of the patients (e.g. age, gender). Decoded personal data will never be shared.

When will the data be made available?

- Upon publication of research results

Data will be made available immediately unless specific IP protections remain to be set.

Which data usage licenses are you going to provide?

If none, please explain why.

The main output of the project will be original scientific research papers. These will adhere to KU Leuven's and FWO's Open Accesss policy. In the context of Open and accessible science, original datasets will be made available with publication, either as supplementary files or using a datasharing platforms such as figshare or Znodo using a CC-BY licence. Upon reasonable and specific request, any data subset and analysis can be made available. For data transfer filesharing via KU Belnet transfer (secure) will be used.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

What are the expected costs for data sharing? How will these costs be covered?

No costs are expected. If any occur, they will be covered by the requesting parties.

Responsibilities

Who will manage data documentation and metadata during the research project?

The PhD and postdoctoral researchers associated with this project (Seyed Ali Elahi- Rocio Castro Vinuelas) will be responsible for data documentation & metadata, under supervision of the PI (Ilse Jonkers).

Who will manage data storage and backup during the research project?

Data management, storage and back up will be performed by the PhD and postdoctoral researchers associated with this project, under supervision of the PI (Ilse Jonkers).

Who will manage data preservation and sharing?

The PI (Ilse Jonkers) will ensure data preservation and sharing.

Who will update and implement this DMP?

The PI(Ilse Jonkers) bears the end responsibility of updating & implementing this DMP.