FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Karel Allegaert, 0000-0001-9921-5105**
Contributor name(s) (+ ORCID) & roles	**Anne Smits, 0000-0002-0710-6698** **Pieter Annaert, 0000-0003-3525-7351**
Project number [1] & title	**Repurposing Paracetamol to ImpRove NEOnatal asphyxia (REPAIR-NEO study), CELSA/24/022**
Funder(s) GrantID [2]	Not applicable
Affiliation(s)	KU Leuven (internal fund, CELSA)

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	Paracetamol is a commonly prescribed drug in newborns to treat pain, fever, or induce closure of the patent ductus arteriosus. Preclinical studies also suggest possible neuroprotective effects in perinatal rodents. This holds the promise of potential repurposing for other indications in perinatal medicine, like moderate to severe hypoxic-ischemic encephalopathy (HIE). Therapeutic hypothermia (TH) is currently the only therapeutic approach with a proven neuroprotective effect (number needed to treat 7). However, mortality or major morbidity is still too present (in 45% of TH cases, instead of 55%). We therefore intend to explore the potential add-on value of paracetamol in TH cases, and develop a preclinical workflow to do so. For these purposes, a modified in vivo model of Rice-Vannucci rat pup neonatal HIE followed by hypothermia and paracetamol treatment will be developed, with in vivo non-invasive MRI/MRS techniques, neurodevelopmental and cognitive-behavioral testing, and ex vivo histological imaging as exposure/effect outcome variables (Charles University) (Obj 1). To further enable translation to the human newborn, this will be supported by in vitro and in vivo data on intravenous paracetamol (Obj 2). The data of Obj 2 will subsequently be integrated in a physiologically-based pharmacokinetic model, specifically developed for TH-asphyxia neonates (Obj 3, I-PREDICT model, KU Leuven). This project may have compound specific impact in the event of positive signals, and evolve to a repurposing drug development plan and trial. Alternatively, we have created a interuniversity preclinical workflow that can be further developed and explored. Irrespective of the results, the collaboration can subsequently be converted to a broader international collaboration between both universities and beyond. When related to data management, KU Leuven is responsible for Obj 2 and Obj 3.

2. 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 [3].

				Only for digital data	*Only for digital data*	*Only for digital data*	*Only for physical data*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
Objective 2 In vitro	In vitro hepatocyte studies, Paracetamol metabolism	☒ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	excel	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Objective 2 In vivo	Pharmacokinetics of iv paracetamol in neonates (in vivo)	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	Excel, RedCap	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Objective 3 PBPK	Application of the iPREDICT PBPK tool.	☒ Generate new data ☐ Reuse existing	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound	xlsx, .csv, .txt, .doc, .pdf, .R, simulation	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB	

[3] Add rows for each dataset you want to describe.

	Quantitative simulations (data-driven predictions) and logs generated following iterative steps taken, and assumptions made, during PBPK model development	data		☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	software specific formats (e.g. Simcyp workspace, compound, population files)	☐ < 5 TB ☐ > 5 TB ☐ NA	

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

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.	**Obj 2, *in vitro*** The reuse of data relates to already existing in vitro drug metabolism data (E4), that also undergo glucuronidation (similar to paracetamol). These in vitro data are owned by a 3rd party (Mithra), and a DTA has been secured (2023-06-09 Neuralis KUL-DTA). The data file and study report is received and stored within KU Leuven. **Obj 2, *in vivo*** The reuse of data relates to already existing PK dataset on paracetamol in neonates, generated by and therefore accessible to the UZ/KU research group (PMID 21317433), while another relevant dataset with data in neonates undergoing TH has been identified (PARASHUTE trial, Sundell Haslund-Krog, PMID 37353311). For this dataset, a DTA has been secured (related to S64850, version 5, signed October 2023), and data access is secured by RedCap.

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 EC approval number: S51597, S64850 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No **Related to Obj 2, *in vitro*** Personal data mentioned on the Product characterisation sheet of "human neonatal & adult hepatocytes": gender, age, race, cause of death, BMI, smoke, alcohol, substance abuse, medical history, infectious diseases. Informed written (parental) consent forms are provided and obtaining them is in agreement with the General Data Protection Regulation (EU) (2016/679) ("GDPR"). **Related to Obj 2, *in vivo*** there will be no new clinical studies, but re-use of already existing data set within an new framework (PBPK, exploration of target attainment exposure) related to clinical trials (S64850, S51597) or DTA transfer agreements (related to S64850). Ethical issues therefore relate to data handling. Short description of the kind of personal data that will be used: name, birth date, birth weight, current weight, gestational age, postnatal age, postmenstrual age, diagnosis on admission, co-medication of included neonates. Data will be pseudonymised and DTA procedures are present.

Will you process personal data[4]? 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) S51597, S64850 ☐ No Additional information: Data will be pseudonymised and DTA procedures are present. **Related to Obj 2, *in vitro*** Personal data mentioned on the Product characterisation sheet of "human neonatal & adult hepatocytes": gender, age, race, cause of death, BMI, smoke, alcohol, substance abuse, medical history, infectious diseases. Informed written (parental) consent forms are provided and obtaining them is in agreement with the General Data Protection Regulation (EU) (2016/679) ("GDPR"). **Related to Obj 2, *in vivo*** there will be no new clinical studies, but re-use of already existing data set within an new framework (PBPK, exploration of target attainment exposure) related to clinical trials (S64850, S51597) or DTA transfer agreements (related to S64850). Ethical issues therefore relate to data handling. Short description of the kind of personal data that will be used: name, birth date, birth weight, current weight, gestational age, postnatal age, postmenstrual age, diagnosis on admission, co-medication of included neonates. Data will be pseudonymised and DTA procedures are present.
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 ☒ No If yes, please comment:
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☒ Yes ☐ No If yes, please explain: The DTA agreements (to UZ Leuven and KU Leuven respectively) rightfully restrict the exploitation of the data to the research topic identified and described. Along the same line, there are agreements on timelines and procedures once data have been used to generate new (scientific) output and results.

[4] See Glossary Flemish Standard Data Management Plan

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 to what data they relate and which restrictions will be asserted.	☒ Yes ☐ No If yes, please explain: The DTA agreements (to UZ Leuven and KU Leuven respectively) rightfully restrict and describe the exploitation of the data to the research topic identified and described. Along the same line, there are agreements on timelines and procedures once data have been used to generate new (scientific) output and results. Legal issues have been handled by KU Leuven, KU LRD and UZ Leuven legal services and advice, and have resulted in the standing DTA for both the in vivo and the in vitro data.

3. 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). *RDM guidance on documentation and metadata*.	**Obj 2, *in vitro*** Documentation will consist of notes in the laboratory notebook (with an electronic backup stored on Box sharing system), that refer to specific experiments and datasets. These notes will describe and document the experimental set-up, biological samples used (certificates of in vitro systems), protocols used, and the names of the respective databases. Initials of the person who conducted the experiments will be noted.? In vivo human documentation: Research methods and practices (including sampling material, volumes, timing, bioanalytical methods used etc) will be documented in the study protocols as word files (EC templates), and additional study SOPs (Standard Operating Procedures) where needed. A blank copy of the informed consent form will also be stored. Final versions of algorithms and scripts will be implemented in manuscripts / research papers, and may be provided in open source platforms e.g. Github (http://github.com/Open-Systems-Pharmacology). All databases will contain definitions of variables and units in a legend section. **Obj 2, *in vivo*** There will be no new clinical studies, but re-use of already existing data set within an new framework (PBPK, exploration of target attainment exposure) related to clinical trials (S64850, S51597) or DTA transfer agreements (related to S64850). As this relates to re-use, these data are already understandable and usable, including e.g. legends, and this will remain similar (taking the DTA agreement into account, as UZ Leuven or KU Leuven are not always the owner of the data). **Obj 3** Quantitative simulations (data-driven predictions) and logs generated following iterative steps taken, and assumptions made, during PBPK model development will be stored. The PBPK model code and source data will be provided upon publication and afterwards in an open source platform (Github (http://github.com/Open-Systems-Pharmacology).

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. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☐ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: For both the in vitro as well as the in vivo data, no metadata standards are available. **Obj 3, *in vitro*** For the in vitro data, we will collect the following metadata: In vitro model system: species, type (suspended/cultured hepatocytes, microsomes,S9 fractions), protein content (mg/mL), cell density (million cells/mL), other donor-specific characteristics (if applicable and available, e.g. specific enzyme activities, number of donors pooled,...); Experimental conditions: model drug, incubation time, incubation concentration, sample volume, incubation temperature, extracellular protein content (type/concentration), sampling scheme, number of replicates; Samples generated: type (activity/abundance), volumes, dilutions, processing steps (centrifugation, filtration or precipitation,...); Endpoints: Activities towards model drugs (uptake and/or metabolic rates or clearance values normalised for protein content or number of cells, protein abundances (pmol/mg protein). Final versions of algorithms and scripts will be implemented in manuscripts / research papers, and may be provided in open source platforms e.g. Github (http://github.com/Open-Systems-Pharmacology). All databases will contain definitions of variables and units in a legend section. **Obj 2, *in vivo*** For the in vivo data, we will collect the following metadata: Patient: study ID, gestational age (GA), postnatal age (PNA), postmenstrual age (PMA), birth weight (BW), current weight (CW), diagnosis on admission; Drug: dose, date and time of administration, route and rate of administration, sampling time, concentration in blood. For the available datasets used for model evaluation (no prospective collection); Biochemical data: albuminemia (g/L), creatininemia (mg/dL), total bilirubin (mg/dL), direct bilirubin (mg/dL); Clinical data: respiratory support (yes/no), mechanical ventilation (yes/no), sepsis (no/suspected/confirmed), co-medication.

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☒ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☐ Other: The J-drive is already used for the iPREDICT project, and for consistency, it is very logic to continue to work within this environments. It provides a minimum capacity of 100 Gb. We anticipate this capacity is sufficient to add the REPAIR-Neo to the I-PREDICT project. As backup, this J drive capacity can be expanded if this were needed
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)
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 ☐ No As backup, the J drive capacity can be expanded if this were needed

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	The J-drive is only accessible for members of the study team, and using a personal login / password. The REDCap (DTA related) is only accessible for members of the study team, with a personal login.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	For the costs of the project and subsequent storage (2 years active project, 10 years in total), we expect a costs of about 1000 euro, and this cost is at present secured by the project, with other resources as back up if this were needed.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five 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...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☒ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain) **Obj 2, *in vitro* and Obj 3** For Obj 2 in vitro and Obj 3 we respect the KU Leuven RDM policy. **Obj 2, *in vivo*** As the UZ Leuven owned data are re-used in the current project, these datasets (if UZ Leuven owned) are already preserved in line with the CTC recommendations. In the event of data access based on DTA, we will respect the signed agreements (as UZ/KU Leuven are not the owner of the data).

Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specify): As data will be 'limited' to electronic data, these data will remained stored at the previously mentioned J drive (budget costs already allocated).
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	For the costs of the project and subsequent storage (2 years active project, 10 years in total), we expect a costs of about 1000 euro, and this cost is at present secured by the project, with other resources as back up if this were needed.

6. 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. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☒ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify: ***Obj 2*** We can only express an opinion on the data that are UZ/KU Leuven owned, and will refer requesting parties to the primary owners (in vivo, in vitro) if this were relevant, requesting the specifications of data sharing. New pseudonymized data, generated during the project, will be made available after publication upon request to the PI as contact person, who will involve all co-supervisors. ***Obj 3*** The PBPK model code and source data will be provided upon publication and afterwards in an open source platform (Github (http://github.com/Open-Systems-Pharmacology).
If access is restricted, please specify who will be able to access the data and under what conditions.	Pseudonymized data will be available upon reasonable request, pending DTA agreement if applicable and a predefined protocol.

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.	☒ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ KU Leuven RDR ☐ Other data repository (specify) ☒ Other (specify) Upon request, to respect the restrictions related to DTAs and the legal and regulatory setting, including but not limited to GDPR.
When will the data be made available?	☐ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify) Data access will be facilitate/enabled (pending agreements on how to share) once the scientific output is in the public domain, irrespective of the timelines of the project (as short, 2 years, it is not unlikely that not all work will already be in the public domain at the end of this project).
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☐ CC-BY 4.0 (data) ☒ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify) I refer to the rationale discussed higher**.**

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☐ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☒ No We will explicitly mention the procedure on how to get potential access to the data in any scientific output, including the limitations related to the standing DTA.
What are the expected costs for data sharing? How will these costs be covered?	In the event of additional costs, these costs should be covered by additional funding. We are aware that legal agreements on DTA also have their costs, but these are - at present - not charged.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The principal investigator, Karel Allegaert will be the responsible to manage data documentation and metadata. In the event of unavailability, the co-investigators are entitled to act on behalf of the principal investigator.
Who will manage data storage and backup during the research project?	The principal investigator, Karel Allegaert will be the responsible to manage data storage and backup. In the event of unavailability, the co-investigators are entitled to act on behalf of the principal investigator.
Who will manage data preservation and sharing?	The principal investigator, Karel Allegaert will be the responsible to manage data preservation and sharing. In the event of unavailability, the co-investigators are entitled to act on behalf of the principal investigator.
Who will update and implement this DMP?	The principal investigator, Karel Allegaert. In the event of unavailability, the co-investigators are entitled to act on behalf of the principal investigator.