D8.1 Data Inclusion Protocol

EQIPD – GA 777364
IMI2-2016-09-03

European Quality In Preclinical Data

WP8 – Data Management

Lead contributor	Tom Van de Casteele (19 – Janssen)
	tvdcaste@its.jnj.com
Other contributors	Maarten Loos (14 – Sylics) maarten.loos@sylics.com
	Emily Sena (1 - UEDIN) Emily.Sena@ed.ac.uk
	Karsten Wicke (20 - Abbvie) karsten.wicke@abbvie.com

Due date	30 Oct 2017
Delivery date	30 Oct 2017
Deliverable type	ORDP
Dissemination level	PU

Description of Work	Version	Date
	V1.0	30 Oct 2017

Document History

Version	Date	Description
V0.1	20 Oct 2017	First Draft
V0.2	27 Oct 2017	Comments
V1.0	30 Oct 2017	Final version

Aim

Provide a transparent protocol for unbiased inclusion of data sets for selected preclinical animal models into the data warehouse (DWH), required for DoAs of work packages 2 (WP 2) and 4 (WP 4)* of this IMI2 consortium agreement, further called the EQIPD-DWH.

*Deviation from DoA: The second part of the deliverable description is left out here: “An ontology for the description of in vivo experiments such as these, including quality aspects and working alongside other groups sharing the same ambition, and develop a template for the depositing of such work in open access fora such as figshare. Our data will be deposited using the template devised.” This refers to development of detailed search algorithms for the identification of relevant published data and is implicitly included in the aim described above.

Objectives

The objectives of this protocol are to define data inclusion and exclusion criteria that will be applied:

1. to select publicly available academic studies and historical data provided by EQIPD consortium members across a range of neuroscience disease models and safety studies of commonly used assays

2. to perform an analysis of historical data to define those variables in study design and data analysis that influence outcome in preclinical neuroscience (focus on Alzheimer’s disease) and safety studies (focus on CNS safety) conducted in industry; and to establish whether these are the same variables that influence outcome in academia

Scope

Several preclinical paradigms will be considered as models for evaluating the scientific hypotheses targeted in this consortium agreement (described in EQIPD WP 2 and WP 4 descriptions). These include the open field or locomotion paradigm, Irwin test and EEG paradigms (sleep, EEG, circadian EEG, pharmaco-EEG, …).

For each considered preclinical paradigm, animal species/strain and intervention/test group, the following types of data will be sought for inclusion in the EQIPD-DWH:

- group summary outcome data and experimental metadata available from published literature (WP 2)

- individual outcome data and experimental metadata available from historical source data sets (WP 2)

- individual outcome data and experimental metadata prospectively collected in this consortium (WP 4)

The scope of this protocol includes all three types of data, but data to be prospectively collected (WP 4) will not be reconsidered for inclusion after they have been collected, and will be included in their totality.

Protocol

The proposed protocol for inclusion/exclusion of WP 2 data for the EQIPD-DWH is based on the practical guide described by Vesterinen et al. (2014) for meta-analyis of data from animal studies. It consists of three main steps:

- Formulation of meta-analysis/systematic review hypotheses that will be addressed
- Data collection/selection
- Documenting and reporting of included and excluded data

These steps are detailed out below.

1. Formulation of meta-analysis hypotheses

The following hypotheses will be evaluated by means of meta-analyses/systematic reviews described in WP 2:

- Research quality factors affect between-study heterogeneity in an experimental effect

- Research quality factors affect between-study heterogeneity in an experimental effect, similarly in industry and academia

Here, research quality factors refer to experimental design factors or factors that reflect rigour in experimental conduct.

2. Data selection

For each considered preclinical paradigm, animal species/strain and intervention/test group, the following stepwise process will be followed to select data for inclusion in the EQIPD-DWH (adapted from Vesterinen et al. 2014):

- Determine the search strategy used to identify relevant literature/source data
- Determine criteria for inclusion or exclusion of identified literature/source data into the EQIPD-DWH
- Determine the data that will be extracted/retrieved
- Identify the primary outcome measure of interest, i.e. the original experimental comparison that will be used as a subject for meta-analysis

Subsequent data processing and analysis steps will be described in meta-analysis/systematic review statistical analysis plans/protocols.

2.1. Determine the search strategy used to identify the relevant literature and source data

A search strategy will be applied that is based on the stepwise process described by Leenaars et al. (2012) for the identification of relevant group summary outcome data for studies published in literature. Here it is adapted to also include a search strategy for the identification of studies for which individual outcome data are available. This adapted search strategy is described in the table below:

Step	Group summary outcome data	Individual outcome data
1. Formulate meta-analysis/systematic review research question	- Heterogeneity in experimental effect* depends on research quality	
2. Identify appropriate databases and sources of studies	- Identify literature databases - Select relevant literature databases	- Identify consortium partners that have data available for the considered preclinical

Step	Group summary outcome data	Individual outcome data
	- Check other sources such as reference lists	paradigm, animal species/strain and intervention/test group
3. Transform research question into search strategy	- Design and run a search strategy customised to each literature database (for an example, see Table 2 in Leenaars et al 2012) - Save citations (titles/abstracts) - Document the applied search strategies	- Identify studies that were executed according to sufficiently similar experimental designs (conditions, outcome measures, data structure)
4. Collect search results and remove duplicates	- Combine saved citations of all literature databases into one file and remove citations that appear more than once	- Identify source data sets for which outcomes have been published**
5. Identify potentially relevant papers and source data sets	- Screen title and abstract of the references and identify papers based on potential relevance	

*For a given preclinical paradigm, animal species/strain, intervention/test group and outcome measure
**Both published data and source data will be included in the EQIPD-DWH

2.2. Determine criteria for inclusion or exclusion of identified literature and source data into the EQIPD-DWH

Once relevant literature and source data have been identified, the following inclusion/exclusion criteria will be applied to decide whether or not an identified study will be included:

- Include all literature and source data that can be retrieved for an established preclinical animal model (i.e. a given animal species/strain subject to a preclinical paradigm), regardless of study outcomes and research quality factors
- Studies executed under a not internally approved preclinical animal model protocol (i.e. a protocol not yet approved at the level of the data contributing consortium partner) will not be excluded
- Studies with documented evidence of batch quality issues with the used experimental/biological materials can be excluded
- Study exclusion should not be based on availability of information on research quality factors: studies lacking information on research quality will be included, and will be labeled as a separate category
- Studies should only be excluded if their inclusion would jeopardize the statistical rigour of the meta-analysis/systematic review as a whole: (a) stud(y)ies could be excluded if their inclusion makes the statistical analysis computationally not feasible, e.g. due to absence of a common data structure, or if their inclusion corroborates assumptions of the statistical model that will be used for the meta-analysis/systematic review; these studies will be documented (including outcome measures) in the final analysis reports

2.3. Determine the data that will be extracted/retrieved

For the sets of studies that will be included in the EQIPD-DWH, the set of variables that will be will be extracted/retrieved from literature and source data bases, and included in the EQIPD-DWH, will be derived from the operationalised version of the ARRIVE guidelines for reporting animal research (Table 2 - METHODS in Kilkenny et al. 2010) and will be refined after discussion. These include, but are not limited to, variables that reflect experimental design (controlled and uncontrolled), variables that reflect candidate outcome measures or from which outcome measures will be derived, variables that reflect aspects of scientific validity (incl. experimental conduct), other variables that could affect outcome measures, etc. … .

Examples include:

- Candidate outcome measures:
  - Open field or locomotion test: distance moved, time moved, speed, …

    - Irwin test: touch response, core temperature, signs of pain, activity, arousal state, …
    - Sleep EEG: %REM, latency to sleep, latency to REM, total sleep, absolute and % sleep, absolute and % awake, …
    - …

- Research quality factors:
    - Randomisation: randomised group allocation, randomised placement in the animal facilities, randomised order for testing/analysis, …
    - Blinding during experiment
    - Blinding the assessment of the outcome
    - Complete reporting of outcome for all animals
    - Listing the outcomes to be measured in an ex ante study protocol
    - Asserting the statistical analysis plan in an ex ante study protocol
    - Whether or not the study was preceeded by a sample size/power calculation
    - Whether or not the study was executed according to an internally approved protocol
    - Whether or not the study included negative and/or positive controls
    - Whether or not data were qc-ed (e.g. checked for reporting errors, …)
    - …

- Other explanatory factors:
    - Circadian time of experiment
    - Light intensity and other environmental parameters
    - Species, strain, sex, age, …
    - Housing conditions (temperature, humidity, single/group), enrichment, health status, …
    - …

A full list of these variables will be made available in the meta-analysis/systematic review statistical analysis plans/protocols.

2.4. Identify the primary outcome measure of interest

For each considered preclinical paradigm, animal species/strain and intervention/test group a primary outcome measure will be selected for use in the meta-analyses/systematic reviews. The selection criterion for the primary outcome measure will be driven by meta-analysis/systematic review design considerations, i.e. minimal data requirements to guarantee robustness of the meta-analyses/systematic reviews. The rationale(s) for these decisions will be described in the meta-analysis/systematic review statistical analysis plans/protocols.

3. Documenting and reporting of included and excluded data

Data included in and excluded from the final meta-analyses/systematic reviews will be documented, including reasons for exclusion, and will be schematically represented in a PRISMA flow chart (Moher et al. 2009).

References

Kilkenny C., Browne W.J., Cuthill I.C., Emerson M., Altman D.G. (2010) Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. PLoS Biol 8(6): e1000412. doi:10.1371/journal.pbio.1000412

Leenaars M., Hooijmans C.R., van Veggel N, ter Riet G., Leeflang M., Hooft L., van der Wilt G.J., Tillema A., and Ritskes-Hoitinga M. (2012) A step-by-step guide to systematically identify all relevant animal studies. Laboratory Animals. 46, 24-31.

Moher D., Liberati A., Tetzlaff J., Altman D.G., The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. https://doi.org/10.1371/journal.pmed.1000097

Vesterinen, H.M., Sena, E.S., Egan, K.J., Hirst, T.C., Churolov, L., Currie, G.L., Antonic, A., Howells, D.W., and Macleod, M.R. (2014). Meta-analysis of data from animal studies: a practical guide. Journal of Neuroscience Methods 221, 92-102.