

DR_PIGS_PID2020-113604RB-I00

1. Resumen de los datos

Proporcionar un resumen de los datos

1.A _ Objectives of the Collected Data and their relation to the General Objectives of the Project.

The data will originate from measurements registered of the performance and behavior of animals in farms (experimental and commercial), and from lab analyses of samples taken from these animals for physiological, genotyping and identification of candidate genes and microbial biomarkers associated with growth performance. It will be used in meeting the project's objectives and in conference and peer-reviewed publications.

Experimental data will be collected by the research team in order to meet objectives: 1 - Deciphering the individual response of low efficient piglets during the nursery period in order to identify discrete probability trees that cover the entire causal hierarchy (Exploratory objective); and

2., Tackling variability by using management and feeding strategies that enhance growth of low efficient piglets (Intervention Objective).

The global effect of all variables tested in the 5 trials will be considered. Discriminant Analysis (DA) will be used to decide if the described features discriminate the extreme groups (**Objective 1**). That is, DA is used to check if these factors permit to classify extreme cases into the given groups and explore the allocation rules. Elements of Causal Inference in regression (as graphical models in Pearl, 2009) are then proposed to obtain accurate estimates of the factors' effects. These methods use not only the data alone, but the knowledge of the data generating process and its temporality, to measure the impact of each factor effect on the responses, even in commercial conditions.

Effects of feeding and management strategies on individuals (**Objective 2**), mainly on distribution tails and variability parameters, will be checked by means of quantile and variance comparison tests, respectively.

1.B Data Type & Format

The majority of the data will be in ASCII (American Standard Code for Information Interchange) data files, eg comma separated variable (CSV) format, which can be imported into rich-text files for word-processing or into spreadsheets. If specialised software is used, then information about free readers will be provided. Data will be generated in the following formats:

Graphics: jpeg, odg, pdf, png, pttx

Tables: odsu, opj, xlsx

Text: docx, pdf, txt

Other: nb, cpp

1.C Re-use of existing data

Not expected

1. D Data Origin

The data will originate from experiments performed along the project at experimental and commercial facilities. Data may also originate from the collaboration with other colleagues at SNIbA

1. E Data Size

The expected size of the data is not currently known, but it is likely to be <10 GB with individual files being ≤50 MB.

1. F Data Suitability

The data will be suitable for use by other research groups working on the following topics: animal efficiency, microbiology, genotyping, animal behavior, nutrition

2. Datos FAIR

2.1 Datos encontrables (incluyendo los metadatos)

2.1.A

The institutional repository provides a unique URL to access the document with the format doi asignado y/o dominio uab [ddd.uab.cat/record/numregistro]

2.1. B

Files will be structured in terms of project and objective identification, and figure and filenames.

2.1. C

All open project results deposited in a repository will provide search keywords together with their metadata. Keywords for open data will be selected from controlled vocabularies that are suitable for the specific type of data.

2.1. D

Version control mechanisms should be established and documented before any data are collected or generated.

2.1. E

The data are expected to be provided in ANSI SQL, XML or text (ASCII) format. For this dataset, data citation and metadata practices derived from the community will be considered.

2.2 Datos accesibles

2.2.A About which data will be made openly available

When no embargo period applies and a data package related to a case study has been marked as public, it will be made openly available. Only data gathered by partners outside of the project work plan and protected by IPR, or inside the work plan but containing confidential information will be kept closed for privacy reasons.

2.2. B About how the data will be made available

Potential users will find out about the data through publications. Data will be made available on publication of the associated paper and will be made accessible on request, under conditions agreed on a case-by-case basis, and after agreement of the project consortium.

Once processing, quality control, organisation, analysis and publication are complete, the data will be made accessible by deposition in open access repositories

2.2. C About methods and software tools needed to access the data

The data will be accessible using the following software: MS Office, Open Office, Adobe Reader, Image Viewer.

2.2. D and E Specific software

Standard publicly available software will be used where possible, but if specialist software tools are developed, i.e. created within Matlab, a short text file (e.g. ASCII) will be provided with the data file to explain the software required.

The majority of the software programmes are available as commercial products or as freeware.

2.2. F Where to store project data

Data, metadata, documents will be stored in the institutional research data repository **DDD**

2.2. G Have you explored the appropriate arrangements with the repository

The [DDD](#), the UAB Digital Repository of Documents, is the tool that collects, manages, disseminates and preserves the scientific, educational and institutional production of the University. The DDD repository shows an organized, open access and interoperable collection.

2.2. H Any restrictions on use?

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

2.2. I Need for a data access committee.

Because of the small scale of these experiments, there is no need for a data access committee.

2.2. J Conditions for access

Free but under conditions established in 2.2.H

2.2. K Identity of the person accessing the data

Identity of the person accessing the data will not be directly ascertained. However, we expect users to follow the standard norms of scientific citation and use of the data in this context will be tracked through scientific citation.

2.3 Datos interoperables

2.3. A Interoperability of our data

The data produced in the project will be interoperable as the datasets will adhere to standardised formats: ASCII, txt, csv, xml, tiff. If MS Office, pdf viewer or image viewer cannot be used, a text (ASCII) file will be provided with the dataset that explains where a free reader can be obtained.

Other types of data have been registered following internal codifications, clearly specified within the file.

2.4 Datos reutilizables

2.4. A How

We will use Creative Commons licences. When reusable dataset is deposited to the repository, an Attribution-NonCommercial license (by-nc) will be requested.

2.4. B When

The data will remain re-usable after the end of the project by anyone interested in it, with no access or time restrictions.

2.4. C Specify whether the data will be usable by third parties and conditions

Each archived data set will have its own permanent repository ID and will be easily accessible. We expect most of the data generated to be made available without restrictions and only data sets subject to IPR and confidentiality issues will be restricted. Where this is going to be the case, agreements will be made based on the individual data sets. Requests for the use of the data by externals will be approved by the project consortium.

2.4. D

The data quality is ensured by different measures. These include validation of the sample, replication and comparison with results of similar studies and control of systematic distortion.

2.4.E

For re-usability the data will be stored on the webpage or on a repository system when implemented for at least ten years.

3. Asignación de recursos

Explicar la asignación de recursos

There are no costs associated to the described mechanisms to make the database FAIR and long term preserved.

The project coordinators or IPs have the ultimate responsibility for the data management in the project

4. Seguridad de los datos

Abordar la recuperación de datos, así como el almacenamiento seguro y la transferencia de datos confidenciales

Institutional servers and backups of the UAB Information Service.

The DDD preservation policy aims to ease and guarantee access, in the medium and long term, to its digital content. The Universitat Autònoma de Barcelona is committed to devote efforts and allocate resources in order to preserve University documents created or managed in a digital support and archived in the institutional repository, to guarantee access to them in the future, and entrusts the management of this goal to the Library Service and the IT Service.

5. Aspectos éticos

Explicar los problemas éticos o legales que pueden afectar la recogida e intercambio de datos

ETHICAL ISSUES: The use of animals for research purposes is necessary in the two main objectives. Due to the lack of alternative procedures available to replace the use of animals, we have designed all the experiments according to European and Spanish laws on the protection of animals used for scientific purposes (Directive 2010/63/EU, R.D. 53/2013) and the 3Rs criteria.

6. Otros temas

Referenciar otros procedimientos nacionales/sectoriales/institucionales/de

financiación usados en la gestión de datos de investigación

As part of Universitat Autònoma de Barcelona commitment to ensuring FAIR and Open data, all research active staff (Postdoctoral fellows, PhD students) are expected to prepare DMPs for their own data, as per the University's Research Data Management Policy. The UAB data management policy defines research data as "the evidence that underpins the answer to the research question and can be used to validate findings regardless of its form." Thus, data covers quantitative and qualitative statements, raw data from measurements and derived data—either cleaned or extracted from a researcher's primary dataset or derived from an existing source.

7. Otro soporte en el desarrollo del plan

Explicar los recursos utilizados para el desarrollo del plan

This DMP has been created with the tool "eiNa DMP" (<https://dmp.csuc.cat>).