DASH-IN: the infrastructure of ENPADASI, using the FAIR concept in nutrition data

Jildau Bouwman, TNO, The Netherlands
Objective

• To deliver:
  • Open access research infrastructure
  • Containing data from a wide variety of nutritional studies
  • Including mechanistic/interventions and epidemiological studies
  • With a multitude of phenotypic outcomes
  • To facilitate combined analyses in the future
ENPADASI Project

- 9 countries
- 51 partners (in 15 national consortia)
- Initial national selection of partners
- Consortium formation in Rome May 2014
- Started December 2014
- 2-year project
- Nationally funded
Why?

- Connect similar studies to resolve chronic diseases with lifestyle related solutions
- Combining studies will increase the power thereby limiting the needs for new and larger intervention studies
- Validation of study results in a different cohort/study will enhance the biological applicability of the conclusions
- Data comparison is instrumental to improve the interpretation and validation of results
- Increase the knowledge and understanding of how food and nutrition can improve human health
Data sharing in Nutrition (DASH-IN)

- Intervention studies (design only)
- Intervention studies (incl clean data)
- Observational studies (incl clean data)
- Observational studies (design only)

Is there a study with similar design? → improve power
Is there a study with similar outcome? → validate outcome
Is there a study with similar design? → Extend measures
Are there studies with specific interventions and outcome
Goal

Federated database

Intervention study

Phenotype database

Observational study

Same wording and grammar

Test case
Current status

• Collect nutritional study data (WP2)
  ➢ Being uploaded

• Standardize terminology (WP4)
  ➢ Mapped, developing new ontologies

• Guidelines for ethics, IP, protection, sharing (WP5)
  ➢ Successful workshop, country specific details being collected, collaboration with CORBEL

• Develop the system to integrate nutritional studies (WP3)
  ➢ Hackathon 8&9th of June: connect Phenotype database and observational studies

• Test the integration in a study cases (WP4)
  ➢ Test case discussion started

• Train users (WP6)
  ➢ Successful workshop, Training material is growing (http://www.enpadasi.eu/wp6.html)
Studies collected

• 23 observational studies

• 12 experimental studies that can be shared and are uploaded

• 21 studies uploaded that are available for data integration
API FAIR compliant

**dataShield Rserver**
modified by ENPADASI
(value-added)

Summary analyses and information
Observational + Intervention studies

Harmonization / OPAL schemes

Metadata MICA / MOLGENIS?

Data in local instances

Intervention studies

Observational studies

Hackathon 8th & 9th of June to develop

*Existing*

*To be analyzed and developed*
<table>
<thead>
<tr>
<th>R</th>
<th>Sampling event</th>
<th>Body fluid</th>
<th>sampling method for body fluid</th>
<th>Study excretion sample</th>
<th>Sampling method</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Exercise</td>
<td>Whole blood</td>
<td>Direct sampling</td>
<td>Urine</td>
<td>Direct collection</td>
</tr>
<tr>
<td>4</td>
<td>Meal</td>
<td>Plasma</td>
<td>Needle</td>
<td>Feces</td>
<td>Swap (absorbent)</td>
</tr>
<tr>
<td>5</td>
<td>Drink</td>
<td>Serum</td>
<td>Swap</td>
<td>Sweat</td>
<td>Other</td>
</tr>
<tr>
<td>6</td>
<td>Diet</td>
<td>White blood cells</td>
<td>Scraping</td>
<td>Menstruation blood</td>
<td>Not defined</td>
</tr>
<tr>
<td>7</td>
<td>Fasting</td>
<td>Red blood cells</td>
<td>Biopsy</td>
<td>Sebaceous matter</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Glucose tolerance test</td>
<td>Saliva</td>
<td>Fractionation</td>
<td>Breath</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Insulin tolerance test</td>
<td>Serum</td>
<td>Other</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Sleep</td>
<td>Lymph</td>
<td>Not defined</td>
<td>Not defined</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Psychological</td>
<td>Cerebrospinal fluid</td>
<td>Not defined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Vaccination</td>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ontology workshop**
The WP4 group proposes the creation of a team of curators in order to:

Create new Ontology entries for ENPADASI based on the work of WP2

<table>
<thead>
<tr>
<th>Randomisation</th>
<th>ORCS_0000058</th>
<th>ORCS:Ontology of Biological and Clinical Statistics</th>
<th>Randomization sampling design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomisation method (if any)</td>
<td>This terms could be new ontologies of ENPADASI.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding (yes/no)</td>
<td>This terms could be new ontologies of ENPADASI.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding method</td>
<td>This terms could be new ontologies of ENPADASI.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study terminated?</td>
<td>This terms could be new ontologies of ENPADASI.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments</td>
<td>EFO_0000727</td>
<td>EFO: Experimental Factor Ontology</td>
<td>treatment</td>
</tr>
<tr>
<td>Number and types of foods</td>
<td>if possible to include lists of foods, drugs or other treatments in order to look for them on ontologic database.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drugs</td>
<td>if possible to include lists of foods, drugs or other treatments in order to look for them on ontologic database.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other treatments</td>
<td>if possible to include lists of foods, drugs or other treatments in order to look for them on ontologic database.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compliance assessed?</td>
<td>This terms could be new ontologies of ENPADASI.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compliance assessment (for each treatment)</td>
<td>This terms could be new ontologies of ENPADASI.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of factors (for factorial designs only)</td>
<td>SIO_000366</td>
<td>SIO:Semanticscience Integrated Ontology</td>
<td>number</td>
</tr>
<tr>
<td>Number of arms (Give number of distinct treatments in study)</td>
<td>SIO_000366</td>
<td>SIO:Semanticscience Integrated Ontology</td>
<td>number</td>
</tr>
<tr>
<td>Number of volunteers enrolled N(Mxx, F: xx)</td>
<td>SIO_000366</td>
<td>SIO:Semanticscience Integrated Ontology</td>
<td>number</td>
</tr>
<tr>
<td>Number of volunteers terminating study N(Mxx, F: xx)</td>
<td>SIO_000366</td>
<td>SIO:Semanticscience Integrated Ontology</td>
<td>number</td>
</tr>
<tr>
<td>Time of recruitment (Start year/end year)</td>
<td>SIO_000391</td>
<td>SIO:Semanticscience Integrated Ontology</td>
<td>time instant</td>
</tr>
<tr>
<td>Follow-up</td>
<td>ICD_000094</td>
<td>ICD:Informed Consent Ontology</td>
<td>follow-up visit</td>
</tr>
<tr>
<td>Describe follow-up</td>
<td>Free text</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Working on a exchange of templates

**Diet intervention (switch)**

Currently, this template contains the following fields. Drag fields to reorder. Drag fields to the list of available fields to remove the field from the template.

- Name (Short text)
- Event name (STRING) (Short text)
- Intervention/Challenge (Dropdown selection of terms)
- Event-type (Dropdown selection of terms)
- Route (Dropdown selection of terms)
- Description (Long text)
- Diet description (File)
- Diet carbohydrate-level (Energy%, Decimal number)
- Diet fat-level (Energy%, Decimal number)
- Diet protein-level (Energy%, Decimal number)

**Available fields**

These fields are available for adding to the template. Drag a field to the template to add it.

- Znd Event-type (Dropdown selection of terms)
- Compound (DBid name) (Dropdown selection of terms)
- Compound dose (Decimal number)
- Compound dose unit (Dropdown selection of terms)
- Compound frequency (Short text)
- Compound full name (Dropdown selection of terms)
- Vehicle (Dropdown selection of terms)

Create new field
Tutorials

Here you will find the tutorials on uploading studies to the ENPADASI Phenotype Database platform.

Intervention study

A Guide to Uploading your Intervention Study to the ENPADASI Phenotype database Platform

Video
DASHIN TUTORIAL VIDEO
Please right-click and save the video to your harddrive.
Note: The video does not work in Windows Media Player, please use Quicktime or VLC.

Descriptor files
ANY STUDY TEMPLATE-intervention file Temporary PLATFORM DETAILS File enpadasi-intervention study
Connection to (ESF)Ris

WP1: Sustainability

WP2: Collect studies

WP3: Federated database
   - Interventional
   - Observational

WP4: Ontologies
   - Integration

WP5: Legal and ethics

WP6: Training
Context & sustainability

WP1  Sustainability
WP2  Collect studies
WP3  Federated database
   - Interventional
   - Observational
WP4  - Ontologies
     - Integration
WP5  Legal and ethics
WP6  Training
Ultimate goal

Biomaterial and data collections

Design of Multi-disciplinary Experiments

International reference data

All resources to measure
- e.g. genomics, transcriptomics, proteomics, metabolomics, bioimaging, microscopy, quantified self, lifestyle, nutritional studies

Medical

Consumer

Improve public health

Model systems

Information & Insight

e-health & quantified self data

All resources for data stewardship and analytics:
- e.g. bioinformatics, informatics, biostatistics, computational (systems) biology, e-science, ICT, …
Risks for ENPADASI

- Continuation of the system after 2016 (maintenance and training)
- Parallel initiatives: thereby not being able to use all shared data at once
Acknowledgements

• All ENPADASI partners

• WP2 – Study data (Lars Dragsted)

• WP3 – Database (Corrado Priami)

• WP4 – Integration (Duccio Cavalieri)

• WP5 – Regulation (Martine Laville)

• WP6 – Training (Lorraine Brennan)

See also: test.dbnp.org (user=user password=useR123! )
MEDICAL DATA BASE OF USEFUL INFORMATION, 5K.
MEDICAL DATA BASE OF USELESS INFORMATION, 500,000,000 MB.