Blueprint for further development of EHES data transfer, management, quality assessment and reporting systems

Prepared by: Ari Haukijärvi¹, Tarja Palosaari¹ and Kari Kuulasmaa¹

¹ Department of Public Health Solutions, National Institute for Health and Welfare (THL), Helsinki, Finland
Disclaimer

The views expressed here are those of the authors and they do not represent the Commission’s official position.

Acknowledgement

The work on this report has been supported by the funding from the EC/DG SANTÉ through BRIDGE Health project (no. 664691).

Cover illustration edited by Hanna Tolonen from vector graphics by Freepik.com.

October 2017 – Helsinki, Finland
**Table of Contents**

Introduction....................................................................................................................................................... 1

Part 1. Data transfer and management system ................................................................................................ 2

Requirements and outline ................................................................................................................................... 2

Checking of the data........................................................................................................................................ 2

Data transfer, import and storage .................................................................................................................. 2

Generating reports ........................................................................................................................................ 2

Future development ........................................................................................................................................ 3

Survey procedures web questionnaire ........................................................................................................ 3

Data upload website and data checking application .................................................................................... 4

Survey database ............................................................................................................................................. 6

Derived variable and indicator database ...................................................................................................... 8

Appendices ..................................................................................................................................................... 10

Appendix 1: Examples of the XML variable declarations and generated outputs ..................................... 10

Appendix 2: ER diagram of the sampling data ............................................................................................ 12

Part 2. Data evaluation and reporting system ................................................................................................. 13

Requirements and outline ............................................................................................................................... 13

Target groups.................................................................................................................................................. 13

What to report................................................................................................................................................ 14

Data evaluation system ................................................................................................................................ 15

Reporting system ........................................................................................................................................... 15

Future development ....................................................................................................................................... 21

Definition of health indicators........................................................................................................................ 22

Definition and calculation of derived variables............................................................................................ 23

Generation of automated data evaluation statistics..................................................................................... 23

Programme codes to calculate health indicators........................................................................................... 24

Metadata repository ....................................................................................................................................... 25
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

Introduction

The European Health Examination Survey (EHES) is a collaboration between organizers of national health examination surveys (HES) in Europe. It supports capacity building in the EU Member States and aims to ensure high quality and comparability of the surveys. It was set up by the EHES Pilot Project in 2009-2012. The European level activities are coordinated by the EHES Coordinating Centre (EHES CC, first called the EHEC Reference Centre), including the external data quality assessment and joint reporting of the survey reports.

A study on the feasibility of conducting comparable HESs in all EU member states ended in year 2008 concluded that, for a proper assessment of the success of the standardization, it is essential that individual level data are collected in a central data repository. This also facilitates rapid reporting and interpretation of the results for European level and easier sharing of the data with research groups for deeper analysis. Therefore, the EHES Pilot Project prepared basic concepts for the data transfer, management, quality assessment and reporting. The EHES CC set up systems for these so that the concepts could be tested using the Pilot Project data from twelve countries.

After the Pilot Project, EHES CC continued some of its tasks as a part of EU Project “Bridging Information and Data Generation for Evidence-based Health policy and research” (BRIDGE Health) in 2015-2017. This report was prepared within the BRIDGE Health project.

In this report, we

• describe the current status of the development of the data transfer, management, quality assessment and reporting system;
• outline the “target” system which would serve optimally the needs of a regular system of national HESs in Europe; and
• give work plans to reach the target for each component of the system.

The report has three parts, where

• Part 1 considers the data transfer and management system;
• Part 2 considers the data evaluation and reporting system; and
• Part 3 summarizes the resources needed for developing and maintenance of the systems and gives a tentative timetable.

1 http://www.ehes.info
4 http://www.bridge-health.eu/
Part 1. Data transfer and management system

Requirements and outline
Establishing a data transfer and management system to facilitate joint collaboration was one of the tasks of the EHES Coordinating Centre (EHES CC) in the EHES Pilot Project. Pseudonymised individual level data were transferred from the EHES Pilot countries to the EHES CC for assessment of the quality of the data and the success of the standardization, and for documentation of country-specific characteristics of the data.

The structure of the system and its functionality are described in detail in EHES Manual, Part C, Chapter 2. The data management services at the EHES CC cover the following main areas and use cases, designed for data transfer, storage, evaluation and analysis:

Checking of the data
- Checking of the survey data locally in each country before uploading the data to the EHES CC.
- Checking of the received survey data in the EHES CC.
- Generating reports of the data checks.

Data transfer, import and storage
- Collecting, storing and updating data on the national survey procedures centrally.
- Uploading the survey data to the EHES CC.
- Importing the survey data to the central database.

Generating reports
- Exporting data for analysis and reports.
- Storing data on population level health indicators and data evaluation statistics calculated from the survey data.

The production of the population level indicators and data evaluation statistics as well as the actual data assessment and reporting of the indicators are a part of the data evaluation and reporting system described in Part 2 of this document.

The core functionality of EHES data transfer and management system is depicted in the Figure 1 as use cases of EHES CC data management services and tools.

---

Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

Future development
This section describes the current status, target and work plan for developing the data transfer and data management system for EHES CC. The system consists of the following five technical parts, which are discussed below:

- survey procedures database and web questionnaire
- survey data upload website
- survey data checking software
- survey database
- derived variable and indicator database

Survey procedures web questionnaire

Current status
In the EHES Pilot Project, a web questionnaire was implemented for collecting data on the national survey procedures. Entering data on national survey procedures to the database at EHES CC allowed comparison of the survey procedures between national manuals, the EHES manual and site visit observations. Currently the questionnaire includes the following topics, with a series of questions for each topic:

- Period of the survey
- Fieldwork staff members and training
- Target population and sampling
- Recruitment
- Communication plan and use of mass media
- Data management
- Order of the measurements and timing of the survey
- Questionnaire administration
- Details on height, weight, waist and blood pressure measurements
- Collection of blood samples
- Preparation of plasma and serum samples
- Non-responder data collection
- Quality control
The questionnaire web application was implemented with Java Servlet and Java Server Pages (JSP) technology to create dynamic web contents and HTTPS protocol with Transport Layer Security (TLS) support to allow a secure communication. The database was implemented on Oracle 11g platform. Although the design pattern of the application is modular and extensible, making changes or adding questions or questionnaire modules requires manual updating of the source code and JSP template(s).

**Target**

A more generic approach should be applied to the definition of the questions and data variables in the questionnaire to facilitate the addition of new questions and changing the questionnaire modules.

The relational database management system (RDBMS) platform should be changed from the commercial to a well-established open source platform.

**Work plan**

- The current system will be replaced by a new one, utilizing the Metaform Project of the ICT Department of the National Institute of Health and Welfare (THL). This is discussed in more detail in Section Survey database of this document. The main features of the new implementation will be:
  - The questionnaire topics, questions and variables are defined in Extensible Markup Language (XML);
  - The XML definitions are transformed into HTML layouts and XML metadata for the web questionnaire;
  - HTML layouts and XML metadata are given as input to the web questionnaire service.

  The web questionnaire service is a parallel project of Metaform, intended for generating web questionnaires and storing the questionnaire data into a database.

- The RDBMS platform will be changed from Oracle to PostgreSQL.

**Needed personnel resources:** Constructing the survey procedures database and web questionnaire requires personnel resources with expertise on data management and XML/HTML (approximately 3 months).

**Data upload website and data checking application**

**Current status**

In the EHES Pilot Project the survey data files were uploaded to the EHES CC via a password protected file upload website. The website was implemented using Java Server Pages (JSP) technology to create dynamic web contents and HTTPS protocol with Transport Layer Security (TLS) support to allow secure transfer of the data.

A data checking application was provided for the national HES organizers so that they could check the survey data before uploading them to the EHES CC. The application was designed to comply with the format of data transfer defined in EHES Manual, Part C, Chapter 2, Appendices 2a...2e. The software was implemented in Java programming language and it was deployed by Java Web Start (JWS) technology to a local use over the network.

---

**Target**

Java Web Start was introduced in 2001 to allow Java applications to be launched through browsers via the Java Network Launching Protocol (JNLP). During the past years the security environment has become increasingly challenging. In 2014, with Java 7 update 51, new security changes were introduced to enhance authentication and authorization for JWS (e.g. code signatures from a trusted authority). Nevertheless, due to vulnerabilities affecting Java plugins and strict security policies being enforced in web browsers, security experts recommend users to disable Java in their browser. Although JWS is a practical way to distribute applications, the technology is slowly becoming obsolete and is being replaced by Web-based technologies. This implies that a new technical solution for the data checking procedure is needed. Because the data checking and upload processes are closely linked, their implementation should be considered together.

The data checking criteria (i.e. constraints, value limits for the data variables and cross check rules between the variables) detect two kinds of data values: values which are illegal (errors) and unusual data values. When constraint violations are detected, the system should allow the national HES organizers to approve and comment the unusual values which are not errors.

**Work plan**

- We propose to implement both the data checking and file upload service using HTML5 technologies. HTML5 is used here in the broad meaning referring both to the W3C HTML5 specification[7] and more generally to the related open web-technologies, such as CSS3[8] and JavaScript programming language.

An overview is shown in Figure 2. In this model HTML5 provides a standard way to interact with local files via the File API specification. The client-side JavaScript handles data files selected by the user. The model allows to implement the data checking step either in the front-end (by client-side JavaScript) or in the back-end (by server-side code), i.e. before or after uploading the data to the Coordinating Centre. The communication is implemented as XHR (XmlHttpRequest) receiving a JSON (JavaScript Object Notation) response. HTTPS connection is used to provide secure transfer of the data.

- Client-side user interface will be implemented, allowing users to approve and comment the values and constraint violations reported by the data checking process.

---

1. W3C HTML5, A vocabulary and associated APIs for HTML and XHTML. W3C Recommendation 2014. [https://www.w3.org/TR/html5/](https://www.w3.org/TR/html5/)
2. W3C Cascading Style Sheets home page. [https://www.w3.org/Style/CSS/](https://www.w3.org/Style/CSS/)

![Figure 2. Overview of the new data checking and file upload application](image)
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

- Data from the uploaded files, including metadata on constraint violations, will be imported into the database at the EHES CC.

**Needed personnel resources:** Creating and testing the data upload and checking system requires personnel resources with expertise on health examination surveys (approximately 0.5 months) and software development (approximately 6 months).

**Survey database**

**Current status**

During the EHES Pilot Project, a central survey database was established at EHES CC to store pseudonymised individual level data on sampling, eligibility, questionnaire items and physical measurements.

Technically, the database was implemented on Oracle 11g RDBMS platform. The software to import the data to the central survey database was implemented in Java programming language. The schema of the survey database along with the description of database tables and components is included in EHES Manual, Part C, Chapter 2 (Appendix “Survey Database Schema”).

**Target**

A more generic and more scalable approach than in the EHES Pilot Project should be applied to the survey data variables and data import/export tool to facilitate the maintenance of the system.

The RDBMS platform should be changed from the commercial to a well-established open source platform.

**Work plan**

Generics and scalability of the database regarding the survey data variables and the data import/export software tool can be implemented by utilizing the XML based technology developed for health surveys in the Metaform Project of the ICT Department of the Finnish National Institute of Health and Welfare (THL).

The main features of this solution are:

- All data structures and variables are described in XML based data definition language;
- The XML documents are further transformed to generate various layouts (html, pdf, etc.) and input for tools that e.g. create the database and process the data;
- The implementation includes software for generating questionnaires, creating a database, and manipulating, importing and exporting data;
- The software is implemented in XSLT, Scala and Java languages. It is independent from hardware and operating system and is transferrable.

The general principle of Metaform is shown in Figure 3.

---

Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

![Diagram of Metaform](image)

*Storage = a list of data variables included in database import or export, defined in XML

The Metaform Project can be applied to the development of the EHES survey database in:

- Defining the data variables: HES data variables, data types and constraints and the related metadata are defined in XML. Examples are given in Appendix 1.
- Creating and/or updating survey database objects: database tables and views are created dynamically on the basis of the XML defined data variables and data types.
- Importing and exporting the database data: the software will be able to use dynamic XML lists of the data variables included in database import and/or export.
- Structuring the database: the relational database has one data table plus one metadata table for each survey data variable (Figure 4). These are linked using the subject identification (ID) as the Primary Key.

![Database structure: a data table for each survey data variable](image)

The model improves the scalability of the survey database and data management tool by making it easier to add survey variables later. The data are versioned, i.e. the history of the data is preserved.

- Storing the sampling data: the data variables to describe up to 3-stage sampling (strata, primary sampling units (PSU), domains and selected persons) were defined in the EHES Pilot Project (Figure
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

5) In the relational database, the Metaform model of one data table for each sampling data variable will be applied. An example of this is given in appendix 2, ER diagram of the sampling data.

![ER diagram of the sampling data](image)

Figure 5. Sampling data in EHES data model

The RDBMS platform will be changed from Oracle to PostgreSQL.

**Needed personnel resources:** Constructing the survey database requires personnel resources with expertise on data management, database design and software development (XSLT, Scala and Java languages) (approximately 18 months).

**Derived variable and indicator database**
The derived variable and indicator database stores the values of the population level indicators and their precision statistics calculated from the data. It also stores derived variables and data evaluation statistics calculated from the survey data. Calculation of the derived variables, data evaluation statistics and indicators are described in Part 2 of this document.

**Current status**
A simplified derived variable and indicator database was created in the EHES Pilot Project (then called quality assessment (QA) database). Definitions of the indicator database were updated after the Pilot Project, in 2017, to include also quantiles and standard deviations of the indicators. The usage of the derived variable and indicator database was tested using the EHES Pilot data.

**Target**
Definitions of the tables and variables of the derived variable and indicator database should be up-to-date whenever the indicator list has been updated. The metadata stored in the metadata repository (see Part 2, Section **Metadata repository**) is linked with the health indicators stored in the derived variable and indicator database in such a way that both data sources, as well as the other survey data, can easily be used jointly for data analysis and for feeding information to different reporting platforms.
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

Work plan

- The calculation of derived variable, data evaluation statistics and indicators is described in Sections Definition and calculation of derived variables, Generation of automated data evaluation statistics, and Programme codes to calculate health indicators. The data will be stored in the same database with the survey data (see Section Survey database of this document), but in a separate logical schema. The calculation and database import will be automated as scheduled tasks.
- Key information to establish a link between the metadata stored in the metadata repository and health indicators stored in the derived variable and indicator database will be created.
- The RDBMS platform will be changed from the commercial to a well-established open source platform.

Needed personnel resources: Creating and testing the derived variable and indicator database requires personnel resources with expertise on data management and databases (approximately 1 months).
Appendices

Appendix 1: Examples of the XML variable declarations and generated outputs

Example: variable declarations

```xml
<!-- ehes.xml (xml root document) -->
<?xml version="1.0" encoding="utf-8"?>
<metaform id="ehes">
  <record ref="ehes_record"/>
</metaform>

<!-- ehes_record.xml (xml variable definitions) -->
<?xml version="1.0" encoding="utf-8"?>
<record id="ehes_record">
  ...
  <!-- Arm circumference -->
  <variable id="bp_armc" type="float" min="15" max="50">
    <title lang="en">Arm circumference</title>
    <unit>
      <title lang="en">cm (round to 0.5 cm)</title>
    </unit>
    <tag domain="sql" type="table">bp_armc</tag>
    <tag domain="csv" type="label">bp_armc</tag>
  </variable>
  <!-- Used cuff size -->
  <variable id="bp_cuff" type="factor">
    <title lang="en">Cuff size</title>
    <values>
      <value code="1">small</value>
      <value code="2">medium</value>
      <value code="3">large</value>
      <value code="4">extra large</value>
    </values>
    <tag domain="sql" type="table">bp_cuff</tag>
    <tag domain="csv" type="label">bp_cuff</tag>
  </variable>
  <!-- Systolic blood pressure -->
  <variable id="bp_syst1" type="integer" min="80" max="250">
    <title lang="en">Systolic blood pressure, 1st measurement</title>
    <unit>
      <title lang="en">mmHg</title>
    </unit>
    <tag domain="sql" type="table">bp_syst1</tag>
    <tag domain="csv" type="label">bp_syst1</tag>
  </variable>
  <!-- Diastolic blood pressure -->
  <variable id="bp_diast1" type="integer" min="40" max="140">
    <title lang="en">Diastolic blood pressure, 1st measurement</title>
    <unit>
      <title lang="en">mmHg</title>
    </unit>
    <tag domain="sql" type="table">bp_diast1</tag>
    <tag domain="csv" type="label">bp_diast1</tag>
  </variable>
  ...
</record>
```
Example: Generated xml definition of database tables for variable data and metadata

```xml
<!-- tables.xml -->
<?xml version="1.0" encoding="utf-8"?>

<!-- Table for data -->
<table name="value_bp_armc" read="true" write="true">
    <column kind="value" type="float" name="val" vref="ehes_record.bp_armc" sql-type="FLOAT"/>
    <column sql-attrs="NOT NULL" key="true" kind="version" name="ver" sql-type="INTEGER"/>
    <column kind="id" key="true" sql-attrs="NOT NULL" name="id" sql-type="INTEGER"/>
</table>

<!-- Table for metadata -->
<table name="meta_bp_armc" read="true" write="true" delete="true">
    <column kind="meta" subkind="comment" name="metatxt" sql-type="TEXT" vref="ehes_record.bp_armc"/>
    <column kind="meta" subkind="class" name="cls" sql-type="TEXT" vref="ehes_record.bp_armc"/>
    <column kind="meta" subkind="id" name="metaid" sql-type="TEXT" vref="ehes_record.bp_armc"/>
    <column kind="meta" subkind="flag" sql-attrs="NOT NULL" name="flags" sql-type="INTEGER" vref="ehes_record.bp_armc"/>
    <column sql-attrs="NOT NULL" kind="version" name="ver" sql-type="INTEGER"/>
    <column kind="id" sql-attrs="NOT NULL" name="id" sql-type="INTEGER"/>
</table>
```

Example: Generated xml definitions for database sql

```xml
<!-- sql.xml -->
<?xml version="1.0" encoding="utf-8"?>

<sql>CREATE TABLE ehes.value_bp_armc (val FLOAT, ver INTEGER NOT NULL, id INTEGER NOT NULL,
    PRIMARY KEY (ver, id));</sql>

<sql>CREATE TABLE ehes.meta_bp_armc (metatxt TEXT, cls TEXT, metaid TEXT, flags INTEGER NOT NULL,
    ver INTEGER NOT NULL, id INTEGER NOT NULL);</sql>

<sql>CREATE VIEW ehes.curr_value_bp_armc AS
    SELECT val,ver,id FROM ehes.value_bp_armc
    NATURAL JOIN
        (SELECT MAX(ver) as ver, id FROM ehes.value_bp_armc
            GROUP BY id) as tmp
    ORDER BY id, -ver;
</sql>
```

Appendix 2: ER diagram of the sampling data
Part 2. Data evaluation and reporting system

Joint evaluation of national surveys and reporting of comparative health information requires European level collaboration within the network of national survey organizers and a central coordinator which is independent of the national surveys. EHES Pilot Project, conducted in 2009-2012, prepared guidelines for health examination surveys, including data evaluation and joint reporting.\textsuperscript{10} Data evaluation and reporting system for EHES is outlined in this report (Figure 6).

\textbf{Requirements and outline}
Evaluation of the national characteristics and the quality of the survey data and reporting of the survey results are the last steps of any health examination survey (HES). Survey results of specified health indicators should be available promptly after the survey fieldwork period. Evaluation of the surveys is essential for correct interpretation of the results. The main time-consuming task relate to the data evaluation, collection of metadata and interpretation of the results, which need to be done manually.

\textbf{Target groups}
There is need to target results of the health examination survey to several different user groups:

- Policy makers who plan and decide on health policies;

Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

- Health professionals and non-governmental organizations who plan and implement preventive actions;
- Ordinary citizens who make daily choices which influence their health;
- Researchers who evaluate the health policies and preventive actions, and improve our understanding on the mechanisms of health; and
- Participants of the surveys who are keen to get their own results on physical and clinical measurements.

Thereby, health indicators can be used by anyone with an interest in understanding the health of the population and its variation across countries. The desired level of information varies between target groups from few core health indicators, needed by policy makers, to a detailed set of stratified health indicators and their metadata used by researchers.

The nature of the provision of the survey participants with their own results is very different from the public reporting of the aggregate data for the other target groups. These are confidential individual level data, reported by the national survey organizers as a part of the survey process. The reporting of their own data to the survey participants will not be given more attention in this document.

**What to report**

The definition of population health indicators must be based on foreseen needs of the target groups to get information and the availability of data. Common specification of the calculation of the information from the HES data is essential for the comparability of the reported information.

Starting in the mid-1990s, work was done to define a set of valid and comparable health indicators for the needs of policy development in the EU and its member states. This was done in collaboration with WHO and OECD which also need such data. This led eventually to a set of 88 European Core Health Indicators (ECHI).

At the global level, WHO has recently created the Global Health Observatory (GHO) database for reporting a regularly updated set of core indicators. This includes indicators to monitor the aim to reduce the number of premature deaths from non-communicable diseases by 25% by year 2025 through nine voluntary global targets. For both ECHI and GHO, reliable data for some of the key indicators can be obtained only through HESs.

In addition to these “core” indicators, we propose to report a comprehensive set of indicators and their break-down by age, socioeconomic status and other population groups which are foreseen to be useful for any of the target groups. The health indicators to be used in joint reporting were specified during the EHES

---

13 World Health Organization (WHO): Global Health Observatory (GHO) data. [http://www.who.int/gho/about/en/](http://www.who.int/gho/about/en/)
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

Pilot Project. The aim is to structure and present the indicators in the reporting system(s) in such a way that each target group can find the needed information easily.

Data evaluation system

The evaluation of the national characteristics and quality of the EHES data includes several steps:

- National level evaluation
- European level evaluation
  - Site visits on the survey fieldwork;
  - Evaluation of the national survey procedures, manuals and questionnaires;
  - Quality assessment of the collected data and success of standardization; and
  - Documenting the used methods and achieved survey quality.

The data evaluation system creates the results for the assessment of the quality and country-specific characteristics of the data that are crucial for correct interpretation of the health indicators. The data evaluation system provides important metadata which should accompany the reports and results of the health indicators. One source for the data evaluation is the survey procedures questionnaire, that is described Part 1, Section Survey procedures web questionnaire of this report. We will focus here on an electronic automated data evaluation system for European level data quality assessment. The data evaluation still requires substantial manual work and expert resources for e.g. comparing questionnaires, assessing deviations from the standard EHES procedures, pointing out shortcomings in data quality and describing the limitations of the data. The results of data evaluation are crucial inputs for the metadata repository described below. A comprehensive data evaluation system helps to prevent misinterpretation of the survey data. In addition, the results will help to plan and organize better health examination surveys in the future.

Reporting system

We propose to implement an automated electronic basic reporting system in three parts:

Derived variable and indicator database

Derived variable and indicator database is a highly automated flexible database to which a comprehensive set of indicators is produced routinely from the individual level survey data. It is also the storage for derived variables and automatically produced data evaluation statistics. The derived variable and indicator database is designed in such a way that it is independent from specific reporting platforms, but it can feed indicators easily to different reporting platforms. Technical details of the derived variables and indicator database are described in Part 1, Section Derived variable and indicator database of this report.

Metadata repository

Metadata are documentation that describes data or dataset. It includes all relevant information that characterizes data. Metadata, or part of it, are sometimes called technical documentation or codebook. Metadata on survey data and health indicators should be stored in a structured (or “tagged”) way that allows its easy use and updating throughout the survey process from the planning to the reporting of the results. Access to viewing and/or editing the metadata must be available widely for planners of the surveys,

---

data processors and users of the survey data and results. General guidelines for presenting metadata from surveys are presented elsewhere. The metadata repository should include data on:

- definition, calculation, rationale and data sources for each variable and indicator;
- information on survey characteristics, data collection, participants, fieldwork and quality monitoring;
- possible exceptions to the standardized survey protocol and the conversion rules for derivation of the EHES variable;
- automated data assessment statistics (e.g. response rates, proportion of missing data, variable distributions etc.); and
- data evaluation results and remarks for each survey.

Some existing metadata tools are described in Table 1. Their usefulness for the needs of EHES will need more thorough investigation.

---

Table 1. Existing metadata tools (examples)

<table>
<thead>
<tr>
<th>Metadata tool</th>
<th>Description</th>
<th>Software</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Documentation Initiative (DDI)</td>
<td>A metadata specification (standard) for describing the data produced by surveys; content, collection of data, questionnaires and data file information. Focuses on survey microdata. A common standard for social, behavioural and economic sciences.</td>
<td>Different commercial and freeware tools created using DDI standard. E.g. metadata editors and survey questionnaire designers (see some of these tools listed below).</td>
<td>The DDI specification is written in Extensible Markup Language (XML).</td>
</tr>
<tr>
<td>The Statistical Data and Metadata Exchange (SDMX)</td>
<td>Standards to describe statistical data and metadata and facilitate exchange of it. It offers an information model for metadata. Focuses on aggregate statistics and data (“macrodata”).</td>
<td>Different commercial and freeware tools created using SDMX standard. E.g. editors, registries, and publishers. See the list of available tools elsewhere.</td>
<td>Uses Extensible Markup Language (XML).</td>
</tr>
<tr>
<td>Colectica Designer and Colectica Reader</td>
<td>For documenting data collection, survey specification, and dataset descriptions and for publishing metadata.</td>
<td>Commercial software.</td>
<td>Uses DDI standard.</td>
</tr>
<tr>
<td>Nesstar technology (Nesstar Publisher, Server and Webview)</td>
<td>Editor for the preparation of metadata and for publishing in the Web. Includes analytical tools and importing and exporting datasets.</td>
<td>Nesstar Publisher (Metadata Editor) is a freeware without support.</td>
<td>Uses DDI standard and Dublin Core.</td>
</tr>
<tr>
<td>Microdata Cataloging Tool (NADA)</td>
<td>A web-based cataloguing system that serves as a portal for browsing, searching, comparing, and downloading metadata.</td>
<td>Distributed as a freeware.</td>
<td>Uses DDI standard.</td>
</tr>
<tr>
<td>SledgeHammer</td>
<td>Software for data conversion and metadata production.</td>
<td>Restricted version is freely available. Compatible with several statistical software.</td>
<td>Uses DDI standard.</td>
</tr>
<tr>
<td>DdiEditor</td>
<td>Tool for data processing and composing data processing of survey datasets. Provides management for questionnaire and data items.</td>
<td>Compatible with SPSS statistical software.</td>
<td>Data documentation using DDI standard.</td>
</tr>
</tbody>
</table>

17 https://www.ddialliance.org/
18 https://sdmx.org/
19 https://sdmx.org/?page_id=4500
20 http://www.colectica.com/
21 http://www.nesstar.com/about/about.html
22 http://www.ihsn.org/software/ddi-metadata-editor
23 http://nesstar-demo.nsd.uib.no/webview/
24 http://www.ihsn.org/software/nada
25 http://www.mtna.us/#/products/sledgehammer
26 https://code.google.com/archive/p/ddieditor/
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

<table>
<thead>
<tr>
<th>Metadata tool</th>
<th>Description</th>
<th>Software</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questasy⁷⁷</td>
<td>Web application to manage the documentation and dissemination of data and metadata for panel surveys. See also LISS Panel Data Archive²⁸. Provides management for questionnaire items and variables and concepts, publications, study information, and more.</td>
<td>Available freely for scientific and non-profit organizations. Compatible with SPSS statistical software.</td>
<td>Database structure is based on the DDI standard, built on PHP framework and uses a MySQL database.</td>
</tr>
<tr>
<td>Maelstrom Research²⁹ - Opal, dataSHIELD and Mica software³⁰</td>
<td>For data documentation, harmonization, storage and sharing.</td>
<td>Open source softwares under the GPL3 licence, that are freely available. Compatible with R statistical software.</td>
<td>User account and password needed.</td>
</tr>
<tr>
<td>Isaacus metadata project in National Institute for Health and Welfare (THL)³¹</td>
<td>An ongoing metadata project that develops a national description model and description system for digital materials in Finland.</td>
<td>THL prepares also a model and description for “statistics production” (based on GSBPM model) and developing the reporting system, that have much in common with the EHES evaluation and reporting systems.</td>
<td>Results and tools not available yet.</td>
</tr>
</tbody>
</table>

**Reporting platform**

The reporting platform makes the health indicators available to those who need them. It is ideally interactive and web based. The user can choose the format of the reports (e.g. table, chart or map) and the desired indicators for specific countries, sexes, age groups etc. for comparison. It should be possible to easily link also information from the metadata repository to the reporting platform.

Existing reporting platforms that could potentially be used for reporting health indicators from EHES are presented in Table 2.

---

²⁸ [https://www.dataarchive.lissdata.nl/](https://www.dataarchive.lissdata.nl/)
²⁹ [https://www.maelstrom-research.org/](https://www.maelstrom-research.org/)
Table 2. Existing reporting platforms (examples)

<table>
<thead>
<tr>
<th>Platform</th>
<th>Description</th>
<th>Output</th>
<th>Stratification and selection</th>
<th>Metadata</th>
<th>Software</th>
<th>Shortcomings</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECHI data tool</td>
<td>Graphic and interactive web based application</td>
<td>Tables, graphs (line and bar charts) and maps. Exporting results in the</td>
<td>Sex and wide age groups, socio-economic status and regional level (when available). Multi-</td>
<td>Definition, calculation, data sources, and rationale for each indicator. Metadata presented</td>
<td>JavaScript Programming Language</td>
<td>Sexes, age groups and countries are hard to present for comparison in simple graphs and tables. Presenting standard errors or confidence intervals is not possible.</td>
</tr>
<tr>
<td>provided by European Commission²²</td>
<td></td>
<td>most common file formats.</td>
<td>selection of indicators for different countries and years.</td>
<td>for selected indicator in different tab in the interactive platform.</td>
<td></td>
<td>Sexes and age groups are hard to present for comparison in maps and tables. Presenting standard errors or confidence intervals is not possible.</td>
</tr>
<tr>
<td>GHO³³ provided by WHO [5]</td>
<td>Interactive repository of health statistics for noncommunicable diseases</td>
<td>Tables, maps and charts. Tables can be downloaded as csv or Excel and</td>
<td>Sex and wide age groups (when available). Selection of indicators for different countries</td>
<td>Definition, data sources and method of estimation for each indicator. Metadata presented</td>
<td>Instant Atlas™ Data Visualization and Presentation Software</td>
<td>Sexes and age groups are hard to present for comparison in maps and tables. Presenting standard errors or confidence intervals is not possible.</td>
</tr>
<tr>
<td>OECD.Stat provided by OECD³⁴</td>
<td>Online platform of statistics on health and health care systems</td>
<td>Tables and charts. Tables can be downloaded as csv, Excel, PC-Axis and</td>
<td>Sex, wide age groups and socio-economic status (when available). Selection of indicator for</td>
<td>as text in web page that is not linked straight from the interactive platform.</td>
<td>Chart software NComVA</td>
<td>Sexes, age groups, socio-economic status and countries are hard to present for comparison in maps and tables. Presenting standard errors or confidence intervals is not possible.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XML.</td>
<td>different countries and years.</td>
<td>Definitions, methodology and sources. Metadata presented as linked pdf-files.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

These platforms would be applicable for the EHES reporting at least for policy makers and ordinary citizen. The lack of possibility to report standard error or confidence intervals limits their applicability to reporting for researchers who evaluate the health policies and preventive actions, and improve our understanding on the mechanisms of health. The mechanism for providing EHES data for these reporting platforms from the EHES Coordinating Centre (EHES CC) would need to be negotiated with the platform providers.

Currently, OECD, Eurostat and WHO-Europe have a Joint Data Collection procedure for receiving indicators and statistics directly from the countries. They provide Excel sheet questionnaires which are filled in by a representative of the national statistical office (correspondent) on predefined indicators. Metadata on indicators are provided by filling in the “Definitions, Sources and Methods joint questionnaire” document. Definitions of the indicators are predefined and cannot be changed by the correspondent. OECD collects further statistics on health care data and WHO/Europe to the Health for All Database on mortality, morbidity, smoking etc. using similar procedures. Data input for these indicator questionnaires cannot be highly automated, but it requires partly manual data processing.

The EHES health indicators stored in the derived variable and indicator database should be in a format that allows easy transfer of results to any existing reporting platforms. In addition, Table 3 gives examples of platform tools, which may be considered for reporting parts of EHES indicator results.

Table 3. Existing platform tools for reporting (examples)

<table>
<thead>
<tr>
<th>Platform</th>
<th>Description</th>
<th>Details</th>
<th>Examples</th>
<th>Technical details</th>
</tr>
</thead>
<tbody>
<tr>
<td>PX-Web (and PX-Win)</td>
<td>Softwares for the Windows and Internet environment to present statistical information.</td>
<td>Creating and converting tables, present tables in graphic form, perform simple statistical operations, etc.</td>
<td>Used e.g. by - Statistics Sweden[^36] - Estonian Health Statistics and Health Research Database[^37] - Statistics Finland[^38]</td>
<td>Created a PC-Axis database, i.e. a catalogue structure containing PC-Axis files.</td>
</tr>
<tr>
<td>provided by the Statistics Sweden[^35]</td>
<td>Interactive web application framework for R statistical software[^40]</td>
<td>Customized presentation of the indicators by creating own applications as R script. Highly flexible outputs as interactive tables, graphs, maps etc.</td>
<td>See Shiny User Showcases[^41]</td>
<td>Requires IT and R programming and statistical skills to create own Shiny web server and reporting platform.</td>
</tr>
</tbody>
</table>

[^35]: https://www.scb.se/sv_/PC-Axis/Start/
[^36]: http://www.statistikdatabasen.scb.se/pxweb/en/
[^38]: http://www.stat.fi/index_en.html
[^39]: http://shiny.rstudio.com/
[^41]: https://www.rstudio.com/products/shiny/shiny-user-showcase/
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

indicator results as figures and also techniques to create reports, documents and reporting web applications, such as Shiny App \(^43\), to disseminate health information using R \(^44\).

There are additional reporting platforms at use at the national level. Their benefits and suitability for European level reporting should also be considered. Examples of such platforms are:

- **Finland - Terveytemme** \(^45\) provided by the National Institute of Health and Welfare, Finland (THL). Platform is created using Atlas reports of InstantAtlas™ software that uses XML data and it is designed for reporting regional differences in health and welfare indicators in Finland. Confidence intervals for prevalence indicators can be presented. THL also has User Interface for Database Cubes and Reports (TIKU\(^46\)) that is based on institute’s own Hydra data format and Amor background software. Currently, TIKU presents only data as tables and basic charts. TIKU is not yet complemented by a metadata database.

- **Norway - Norhealth** \(^47\) provided by the Norwegian Institute of Public Health is a web-based health information system that monitors health, including risk- and protective factors. Indicators are presented at the national and regional level for Norway. Confidence intervals or standard errors are not reported. Norhealth uses Nesstar technology in presenting the data (the same technology is used e.g. by European Social Survey). \(^48\)

- **Germany - The Information System of the Federal Health Monitoring** \(^49\) provided by Robert Koch Institute and Federal Statistical Office, Germany, is a data infrastructure which presents comprehensive health information as tables, diagrams and texts. Confidence intervals or standard errors are not reported. It is developed by using Robotron Datenbank-Software GmbH and Oracle Database Online Analytical Processing (OLAP) tool.

- **Netherlands - Volksgezondheidenzorg** \(^50\) provided by National Institute of Public Health and the Environment (RIVM) is the gateway to information about health and disease, risk factors, care and prevention in the Netherlands. Confidence intervals are given in tables for some prevalence indicators.

**Future development**

This section describes the current status, target and work plan for developing the evaluation and reporting system for EHES. The target and the work plan are specified separately for construction and maintaining the EHES evaluation and reporting system. National health examination surveys conducted in 2007-2017 are planned to be included in the construction phase and used for testing the system (total of 31 surveys, Table 4). It is assumed that 6 national surveys per annum are included in the maintaining phase for the EHES data management, evaluation and reporting system.

---

\(^{42}\) [http://www.ehes.info/rc/tools/EHES_reporting_options_using_R_publish.html](http://www.ehes.info/rc/tools/EHES_reporting_options_using_R_publish.html)

\(^{43}\) [https://shiny.rstudio.com/](https://shiny.rstudio.com/)

\(^{44}\) [https://www.r-project.org/](https://www.r-project.org/)

\(^{45}\) [http://www.terveytemme.fi](http://www.terveytemme.fi)


\(^{47}\) [http://www.norgeshelsa.no/norgeshelsa/](http://www.norgeshelsa.no/norgeshelsa/)


\(^{49}\) [http://www.gbe-bund.de](http://www.gbe-bund.de)

\(^{50}\) [https://www.volksgezondheidenzorg.info/](https://www.volksgezondheidenzorg.info/)
Table 4. Health examination surveys in EU Member States in 2007-2017

<table>
<thead>
<tr>
<th>Country</th>
<th>Name of the survey</th>
<th>Year(s) when conducted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Republic</td>
<td>Czech-EHES</td>
<td>2014-2015</td>
</tr>
<tr>
<td>Denmark</td>
<td>The Danish Health Examination Survey (KRAM)</td>
<td>2007-2008</td>
</tr>
<tr>
<td>Finland</td>
<td>FINRISK</td>
<td>2007 and 2012</td>
</tr>
<tr>
<td></td>
<td>Health2011</td>
<td>2011-2012</td>
</tr>
<tr>
<td></td>
<td>FinHealth</td>
<td>2017</td>
</tr>
<tr>
<td>France</td>
<td>Étude Nationale Nutrition Santé (ENNS)</td>
<td>2006-2007</td>
</tr>
<tr>
<td></td>
<td>ESTEBAN</td>
<td>2014-2016</td>
</tr>
<tr>
<td>Germany</td>
<td>Studie zur Gesundheit Erwachsener in Deutschland (DEGS)</td>
<td>2008-2011</td>
</tr>
<tr>
<td>Greece</td>
<td>HYDRIA</td>
<td>2013-2014</td>
</tr>
<tr>
<td></td>
<td>National Morbidity and Risk Factor Study (EMENO)</td>
<td>2014-2015</td>
</tr>
<tr>
<td>Ireland</td>
<td>SLÁN</td>
<td>2007</td>
</tr>
<tr>
<td>Italy</td>
<td>OEC/HES</td>
<td>2008-2012</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>EHES-LUX</td>
<td>2013-2015</td>
</tr>
<tr>
<td></td>
<td>ORISCAV-LUX 1</td>
<td>2007-2008</td>
</tr>
<tr>
<td></td>
<td>ORISCAV-LUX2</td>
<td>2016-2017</td>
</tr>
<tr>
<td>Malta</td>
<td>SAHHTEK</td>
<td>2014-2016</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Nederland de Maat Genomen (NL de Maat)</td>
<td>2009-2010</td>
</tr>
<tr>
<td>Poland</td>
<td>WOBASZ II</td>
<td>2013-2014</td>
</tr>
<tr>
<td>Portugal</td>
<td>Inquérito Nacional de Saúde com Exame Físico (INSEF)</td>
<td>2015-2016</td>
</tr>
<tr>
<td>Slovakia</td>
<td>EHES</td>
<td>2011</td>
</tr>
<tr>
<td>UK/England</td>
<td>Health Survey for England (HSE)</td>
<td>Annually</td>
</tr>
</tbody>
</table>

Definition of health indicators

Current status

A comprehensive set of health indicators for EHES core measurements using comparable data are defined in EHES Manual, Part C. These are in line with the European Core Health Indicators (ECHI).

51 http://www.szu.cz/ehes
52 http://www.niph.dk/Forskning/Sundhedsvaner/KRAM.aspx
57 http://invs.santepubliquefrance.fr/Dossiers-thematiques/Environnement-et-sante/Esteban
58 http://www.rki.de/DE/Content/Gesundheitsmonitoring/Studien/Degs/degs_w1/degs_w1_node.html
60 http://emeno.gr/
61 http://www.ucd.ie/issda/data/surveyonlifestyleandattitudestonutritionslan/
62 http://www.cuore.iss.it/
64 http://www.oriscav.lih.lu/Accueil.aspx
65 http://www.rivm.nl/Onderwerpen/N/Nederland_de_Maat_Genomen
66 http://www.wobasz-projekt.pl/
67 http://www.insef.pt/English/Pages/INSEF.aspx
68 http://content.digital.nhs.uk/healthsurveyengl
Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

**Target**
Indicators are estimates for all foreseeable breakdowns of the population, such as by sex, age and socioeconomic status. Definitions of indicators require regular updating and further development (e.g. new indicators for tests of functional capacity; hand grip test and chair stand test, and for urine samples; sodium intake).

**Work plan**
- Indicators for new measurements specified in the EHES Manual will be defined;
- The rationale and interpretation of each indicator will be described.

**Needed personnel resources:** In the construction phase, defining indicators requires expertise on health examination surveys (approximately 0.5 month) and health information needed by the target user groups and following up the development of ECHI and other health indicator projects.

**Definition and calculation of derived variables**
Derived variables are individual level variables derived from the survey data and survey metadata. They facilitate the calculation of the survey- and population-level evaluation statistics and health indicators.

**Current status**
Derived variables corresponding to the EHES core indicators are defined in EHES Manual, Part C, Chapter 5. 69 During the Pilot Project, the R programme codes were written to calculate the individual level derived variables for the database. The R-programmes were tested in the EHES CC using data from the pilot surveys.

**Target**
Definitions of derived variables require updating whenever new indicators or evaluation statistics are defined.

**Work plan**
- Derived variables for the new indicators will be defined and the programme codes will be modified accordingly.
- Derived variables will be saved to the derived variable and indicator database (see Part 1, Section Derived variable and indicator database).

**Needed personnel resources:** Derivation and proper testing of the derived variables in the construction phase requires personnel resources with expertise on health monitoring (approximately 0.5 month) and statistical computing (approximately 1 month) and full-sized survey data.

**Generation of automated data evaluation statistics**

**Current status**
During the EHES Pilot Project a simple automated evaluation system was outlined, prepared and tested at the EHES CC. The work covered the automated calculation of evaluation statistics and outlines for data

---

Blueprint for future development of EHES data transfer, management, quality assessment and reporting systems

evaluation reports. These are documented in the EHES Manual, Part C, Chapter 3\textsuperscript{70}, but the actual data evaluation reports for the pilot surveys were not prepared. Evaluation statistics were generated using R programme codes and the results were stored in the derived variable and indicator database. Static web pages including the main data evaluation tables were created for being included in data evaluation reports.

**Target**
Data evaluation results, covering survey procedure descriptions, participation rates, proportions of missing and incomplete data, distribution tables and plots for each data item, and measurement-specific quality statistics (such as proportions of terminal digits), are generated routinely to a database and reported for each EHES participating survey.

**Work plan**
- A system to generate evaluation results for the derived variable and indicator database will be developed using R statistical software and suitable metadata tools.
- An automated and interactive platform to present and extract evaluation tables for those who prepare data evaluation reports for the surveys will be created.

**Needed personnel resources:** The development of the automated data evaluation system in the construction phase requires personnel resources with ICT (approximately 1 month) and statistical programming skills (approximately 2 months) and expertise on heath examination surveys and their evaluation (approximately 0.5 month), and proper data management, metadata and reporting tools.

**Programme codes to calculate health indicators**

**Current status**
During the EHES Pilot Project, R programme codes to calculate the indicators were prepared and tested with the pilot survey data. The indicators were calculated separately for men and women by 10-year age group, and age-standardized results for 25-64 year-olds assuming a flat age distribution among the 10-year age groups (equal weights for every age group). Programme codes that take sampling design into account was prepared after the Pilot project, in 2017.

**Target**
Weighting for non-response is needed in the calculation of the indicators to improve the representativeness of the survey for the underlying population. Health indicators need to be stratified for also socio-economic factors.

**Work plan**
- R-programme codes will be amended to take into account non-response weights and stratifications by socio-economic status. The amended programme codes will be tested using full-sized survey data.
- Indicator results will be saved to the derived variable and indicator database.

**Needed personnel resources:** This work requires in the construction phase personnel with statistical computing skills (approximately 4 months) and full-sized survey data for testing.

Metadata repository

Current status
A brief review of tools for the metadata repository which are potentially useful for EHES has been made (see Section Reporting system and Table 1).

Target
Creating and operating a system for storing metadata on EHES data. The system will allow easy updating and retrieval of metadata for reporting and evaluating the indicators.

Work plan
- The review of available standards for metadata specifications and available metadata storing systems will be completed and their suitability for EHES reporting will be assessed.
- Based on the review, suitable tools for creating metadata repository and for creating and editing metadata in the repository will be chosen. If necessary, needed bits of the metadata system will be programmed.
- Relevant metadata will be collected and entered to the metadata repository.

Needed personnel resources: Creating and testing the metadata repository in the construction phase requires personnel resources with data management and data base skills (approximately 3+1 months for setting up and testing), knowledge about metadata standards (approximately 2+1 months for setting up and testing) and substance expertise on health examination survey protocols and data quality (approximately 3 months for setting up and 1.5 months per included survey for testing and evaluating the received data, making total of 52,5 months). The amount of work and expertise needed will depend on how much the meta-data system can be based on existing tools. Here it is assumed that the metadata repository can be set up using existing metadata tools. Nevertheless, creating and keeping metadata information exhaustive and up to date still requires a substantial effort on top of that (see Part 3 of this document).

Reporting platform

Current status
A brief review of existing reporting tools and platforms has been made (see Section Reporting system and Tables 2 and 3).

Target
The reporting platforms read data form the derived variable and indicator database. A wide selection of relevant health indicators from national HESs will be available for different target user groups via major European and international reporting platforms. Where necessary for reporting sufficient details, additional platforms will be implemented.

Work plan
- The requirements and possibilities of existing reporting platforms from the point of view of EHES data will be examined.
- A complementary reporting platform will be developed for data that cannot be reported using the existing platforms (e.g. further stratified results and measures of uncertainty, confidence intervals, standard errors etc.).
**Needed personnel resources:** The review of the specification of the existing platforms requires skills on disseminating health indicators, system development, and data management (approximately 3.5 months in total). Creating and testing a possible complementary reporting platform requires also expertise on statistical programming (approximately 8 months in total).
Part 3. Summary of the needed resources

The person months of different experts needed to do the work identified in Parts 1 and 2 are summarized in Table 5. The total need is 28.5 person months (2.6 full-time equivalent, fte) for constructing and testing the EHES data transfer and management system and 77.5 person months (7.0 fte) for constructing and testing the EHES evaluation and reporting system. The total needed for the two (i.e. 106 person months=9.6 fte) splits to 53 person months (4.8 fte) for substance knowledge, 39 person months (3.5 fte) for ICT and data management, and 14 person months (1.3 fte) for statistical programming. This includes the evaluation of reporting for the all national HESs conducted in 2007-2017 in the EU member states. This will be needed for testing the systems.

Parts 1 and 2 of this document, and hence the preceding paragraph, concern the construction and testing of the systems. Thereafter, resources will be needed for maintaining and operating the systems for future national HESs. Estimates for the annual personnel resources for the maintenance and operation, assuming about 5 national HESs each year, are also given in Table 5. This requires 1.5 fte experts on substance of health examination surveys (i.e. 16.5 person months), 0.8 fte ICT and data manager (i.e. 9 person months) and 0.5 fte statistician (i.e. 6 person months) which makes about 31.5 working months (2.9 fte) in total. The work will consist of support to the national HES organizers on data management and data transfer (6 person months of different experts), data checking and processing in the EHES CC (6 person months), data evaluation (12 person months) and reporting (7.5 person months).

These estimates for the need of personnel resources concern only the maintenance and operation of the data transfer, management, data assessment and reporting at the EHES CC. They do not include other activities of the EHES CC, such as coordination of the EHES Network, management (including data transfer agreements), support to the countries in planning surveys, quality assurance etc. (see the EHES Manual, Part C, Chapter 6.2).

Table 5. Summary of the needed personnel resources in the construction and maintaining phases

<table>
<thead>
<tr>
<th>Expertise</th>
<th>Constructing and testing the data transfer and management system</th>
<th>Constructing and testing the reporting and evaluation system</th>
<th>Maintaining and operating the EHES data management, evaluation and reporting system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substance knowledge in health examination surveys</td>
<td>0.5 person months</td>
<td>52.5 person months</td>
<td>16.5 person months</td>
</tr>
<tr>
<td>ICT and data management</td>
<td>28 person months</td>
<td>11 person months</td>
<td>9 person months</td>
</tr>
<tr>
<td>Statistical programming</td>
<td>14 person months</td>
<td>6 person months</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>28.5 person months</strong></td>
<td><strong>77.5 person months</strong></td>
<td><strong>31.5 person months</strong></td>
</tr>
</tbody>
</table>