# POWER2DM

**"Predictive model-based decision support for diabetes patient empowerment"**

**Research and Innovation Project**
**PHC 28 – 2015: Self-management of health and disease and decision support systems based on predictive computer modelling used by the patient him or herself**

# POWER2DM D4.6

# Data Quality Analysis Framework II (D4.4.1b)

| | |
|---|---|
| *Due Date:* | 31st October 2016 (M21) |
| *Actual Submission Date:* | 31st October 2016 (M21) |
| *Project Dates:* | Project Start Date: February 01, 2016<br>Project End Date: July 31, 2019<br>Project Duration: 42 months |
| *Deliverable Leader:* | PD |

| **Project co-funded by the European Commission within H2020 Programme (20015-2016)** | | |
|---|---|---|
| **Dissemination Level** | | |
| **PU** | Public | X |
| **PP** | Restricted to other programme participants (including the Commission Services) | |
| **RE** | Restricted to a group specified by the consortium (including the Commission Services) | |
| **CO** | Confidential, only for members of the consortium (including the Commission Services) | |

**Document History:**

| Version | Date | Changes | From | Review |
|---|---|---|---|---|
| V0.1 | 08.08.2016 | Initial overview made in Excel for review by TNO. | PD | - |
| V0.2 | 12.08.2016 | Comments and feedback from TNO in Excel. | TNO | PD |
| V0.3 | 04.10.2016 | Further completion by PD and comments and feedback from PD in Excel. | PD | |
| V0.4 | 04.10.2016 | Excel information transferred to project template for upload in ECAS. This deliverable needs to be updated at M10 and will then be finalized. Currently not all data types and quality are known that are needed for the first pilot. | PD | European Commission |
| V0.5 | 14.04.2017 | Updated deliverable based on review comments by European Commission | PD | LUMC, TNO |
| V0.6 | 19.04.2017 | Updated table 2 and 3 based on feedback TNO | PD | |
| V2.0 | 31.10.2017 | Updated chapter 3 and added chapter 4 for final delivery Data Quality Analysis Framework D4.4.1b. Deliverable 4.5 renamed to deliverable 4.6 | PD | |

**Contributors (Benef.)** Marlies Schijf (PD)
Roosmarijn Schopman (PD)
Albert De Graaf (TNO)

**Responsible Author** Marlies Schijf **Email** m.schijf@primedata.nl

# POWER2DM Consortium Partners

| Abbv | Participant Organization Name | Country |
|---|---|---|
| TNO | Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek | Netherlands |
| IDK | Institute of Diabetes “Gerhardt Katsch” Karlsburg | Germany |
| SRDC | SRDC Yazilim Arastirma ve Gelistirme ve Danismanlik Ticaret Limited Sirketi | Turkey |
| LUMC | Leiden University Medical Center | Netherlands |
| SAS | SAS Servicio Andaluz de Salud | Spain |
| SRFG | Salzburg Research Forschungs Gesellschaft | Austria |
| PD | Prime Data B.V. | Netherlands |
| iHealth | iHealth EU | France |

# ROADMAP

| Doc. Version | Due Date | Expected/Planned Actions | From |
|---|---|---|---|
| V0.4 | M8 | Publish the deliverable (DEADLINE) D4.5 (D4.4.1a Data Quality Analysis Framework I). This is a draft version that needs to be updated in M10. | PD, TNO |
| V0.6 | M15 | Finalisation of D4.5 (D4.4.1a Data Quality Analysis Framework I) as a baseline for the first prototype.<br>Adjustments based on comments of intermediate review were made. Sections 3.3 and 3.4 were deleted in favour of adding additional information to the tables of section 3.2, such as extra parameter characteristics and the mapping to the data model of the Personal Data Store. | ALL |
| V2.0 | M21 | Publish the deliverable (DEADLINE) **D4.6** (D4.4.1b Data Quality Analysis Framework II).<br><br>This document will be a reference to a DEM: Demonstrator, pilot, prototype. It is part of Prototype and System Release 2.<br><br>Final version | PD |

# OPEN ISSUES

| No: | Date | Issue | Resolved |
|---|---|---|---|
| ~~1~~ | ~~04.10.2016~~ | ~~Document is not reviewed yet by partners other than TNO. Document is currently based on the quantification protocol, which might not represent the complete scope for the POWER2DM prototypes.~~ | **~~M15~~** |
| ~~2~~ | ~~04.10.2016~~ | ~~Definitions and Acronyms have not been listed yet.~~ | **~~M15~~** |
| ~~3~~ | ~~04.10.2016~~ | ~~Sections 3.3 and 3.4 were deleted in favour of adding additional information to the tables of section 3.2, such as extra parameter characteristics and the mapping to the data model of the Personal Data Store.~~ | **~~M15~~** |
| 4 | 31.10.2017 | For some of the devices and corresponding parameters discussion is still ongoing by and between the clinical partners whether they will be used and if so how they could be integrated into the POWER2DM system consequently. | |

# TABLE OF CONTENTS



# LIST OF TABLES

# 1 PURPOSE

Data measured by devices or gathered through clinical questionnaires needs to be validated for correctness to ensure correct action plans for the patient as well as reliable modelling results for the models KADIS, MARVEL and the risk score calculations. Several quality checks and detections will be implemented in the POWER2DM Data Integration Platform, such as outlier detection and min-max analysis, in order to overcome statistical issues within the datasets. Moreover, basic data processing will take place in order to guarantee data quality of the available data via API's. Both data quality analysis and basic data processing have been made configurable and changeable through configuration.

Originally the quantification campaign is carried out with the existing functionality in PatientCoach. The PatientCoach functionality was the starting point for this deliverable: the Data Quality Analysis Framework. Afterwards, this deliverable has been updated with the adjusted scope for prototype 2, including an indication of which parameters will be in scope for prototype 2 and which are still being considered by the clinical partners for future implementation.

The SMSS mobile application, web interface and portable measuring devices will be used to collect all patient input data, which will be stored in the Personal Data Store (PDS). From there several applications can retrieve the data, for instance to be used as input for modelling or to create overviews and visualisations in user interfaces.

The purpose of this deliverable is to make an overview of data validation and data processing logic that is needed in the POWER2DM Data Integration Platform. This work is done in four steps. First measuring devices and questionnaire questions – both quantitative and qualitative data – are selected per prototype, based on the Quantification Campaign Protocol and general data quality criteria. Secondly, an inventory of all related data types and their characteristics is created. As a third step boundary conditions are determined for these data types, which have been translated to automated checks (business rules) in the POWER2DM Data Integration Platform. Finally, the fourth step makes the link between external data and the POWER2DM Personal Data Model (FHIR).

# 2 REFERENCE DOCUMENTS

The following documents were used or referenced in the development of this document:

- POWER2DM D4.1 Personal Data Model and Service API
- Devices used in the POWER2DM Quantification Campaign
- C.1 Power2DM Quantification Campaign Protocol LUMC v1_9 2016-07-21

The following websites were used or referenced in the development of this document:

- Use and Abuse of HOMA Modeling, American Diabetes Association, Diabetes Care via http://care.diabetesjournals.org/content/27/6/1487
- Mayo Clinic Cholesterol Test Results via http://www.mayoclinic.org/tests-procedures/cholesterol-test/details/results/rsc-20169555
- Understanding Your Lab Values, National Kidney Foundation, via https://www.kidney.org/kidneydisease/understandinglabvalues
- Microalbuminuria from Wikipedia, the free encyclopedia, via https://en.wikipedia.org/wiki/Microalbuminuria
- Creatinine from Wikipedia, the free encyclopedia, via https://nl.wikipedia.org/wiki/Creatinine
- iHealth Wave - Activity, Swim and Sleep tracker, via https://ihealthlabs.eu/en/48-ihealth-wave.html
- iHealth Edge via https://ihealthlabs.com/fitness-devices/ihealth-edge
- The Science Behind Spire, via https://spire.io/science
- Fitbit HR Charge 2 heart rate + fitness wristband, via https://www.fitbit.com/charge2
- Fitbit Charge 2 Product Manual, via https://staticcs.fitbit.com/content/assets/help/manuals/manual_charge_2_en_US.pdf
- Target Heart Rates – American Heart Association, via http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/FitnessBasics/Target-Heart-Rates_UCM_434341_Article.jsp

## 2.1 Definitions and Acronyms

**Table 1 List of Abbreviations and Acronyms**

| Abbreviation/ Acronym | DEFINITION |
|---|---|
| **PDS** | Personal Data Store (database of the POWER2DM system) |
| **AS4** | Axis IV: Psychosocial and environmental factors contributing to the disorder |
| **BMI** | Body Mass Index |
| **HbA1C** | Glycated hemoglobin; a form of hemoglobin that is measured primarily to identify the three-month average plasma glucose concentration. |
| **HDL** | high-density lipoprotein cholesterol |
| **LDL** | low-density lipoprotein cholesterol |
| **QoL** | Quality of Life |
| **SMSS** | Self-Management Support System |
| **T1DM** | Type 1 diabetes mellitus |
| **T2DM** | Type 2 diabetes mellitus |

# 3 DATA QUALITY ANALYSIS FRAMEWORK

## 3.1 Overview

The following section gives an overview of measured and registered data types and their characteristics. These data types are data that are – at first – external to the POWER2DM system. When these data types enter the POWER2DM system, they will need to be validated and mapped to the data types that were defined to be used within the POWER2DM system, see D4.1 Personal Data Model and Service API.

Our methodology is as follows:

1. Identification of measuring devices and questionnaire questions based on use cases and data quality criteria;
2. Definition of relevant parameters and their characteristics;
3. Definition of boundary conditions for these data types;
4. Mapping of these data types to the POWER2DM Personal Data Model.

The tables in the following subsections give a detailed overview of the results of each of the above steps.

## 3.2 Functional Description of Data Types

The key ID in the tables in this section is Parameter, which is listed in the first column. The Parameters have been clustered into 10 categories:

1. Patient Characteristics
2. Clinical Characteristics
3. Blood Glucose Level
4. Eating Behaviour
5. Exercise and Energy Expenditure
6. Sleep
7. Stress
8. Mood
9. Diabetes Medical Treatment
10. Questionnaire: Qualitative information

For each parameter the following characteristics have been inventoried:

1. Measuring device
2. More specific measuring device/method
3. Physical property to measure/register
4. Unit of measurement
5. Valid range lower cut-off
6. Valid range higher cut-off
7. Frequency and Timing of measurement/registration
8. (Expected) Data quality
9. Comments
10. Mapping of data type to Personal Data Store (PDS) data model (FHYR)

**Please note** for the following subsections that:

- Information in black is relevant for prototype 2, which will be used for the feasibility studies;
- Information in grey is still considered by clinical partners whether it needs to be included into the system at some point.

### 3.2.1 Data Types of Patient Characteristics

**Table 2 Parameters and measuring devices for category Patient Characteristics**

| Patient Characteristics | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Reason to start POWER2DM** | Shared decision making application | All patient characteristics will be recorded and input into the patient data entry form at baseline (T0) by the diabetes healthcare provider or researcher. | | | | | | | | |
| **Age** | Shared decision making application | | Age | Years | 10 | 111 | 1x baseline | Good, factual | | Patient |
| **Gender** | Shared decision making application | | Gender | Male or Female | | | 1x baseline | Good, factual | | Patient |
| **Height** | Shared decision making application | | Height | cm | 100 | 220 | 1x baseline | Good, provided it is measured correctly | | SimpleQuantity Observation |
| **Type Diabetes** | Shared decision making application | | Type Diabetes | T1DM or T2DM | | | 1x baseline | Good, provided it is diagnosed correctly | | Condition |
| **Illness History** | Shared decision making application | | Illness History | diverse | | | 1x baseline | Good, provided this is well done by healthcare provider or researcher | | Condition |
| **AS4** | Shared decision making application | Diagnostic and Statistical Manual of Mental Disorders (DSM) | Axis IV: Psychosocial and environmental factors contributing to the disorder | diverse | | | 1x baseline | Good, provided it is communicated truthfully | | |
| **Weight** | Shared decision making application | Weight and BMI (calculated from Weight) will be reassessed at the end of Phase 1 and Phase 2 | Weight | kg | 10 | 250 | 1x baseline, 1x wk 11-12 | Good, provided it is measured correctly | | SimpleQuantity Observation |
| **Waist** | Shared decision making application | | Waist | cm | 40 | 175 | 1x baseline, 1x wk 11-12 | Good, provided it is measured correctly | | SimpleQuantity Observation |
| **Waist-Hip ratio** | Shared decision | | Waist-Hip ratio | | 0.6 | 1.4 | 1x baseline, 1x wk 11-12 | Good, provided it is measured correctly | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | making application | | | | | | | | | |
| **BMI** | Shared decision making application | Weight and BMI (calculated from Weight) will be reassessed at the end of Phase 1 and Phase 2 | BMI | kg/m^2 | 10 | 75 | 1x baseline, 1x wk 11-12 | Good, provided it is calculated correctly | | SimpleQuantity Observation |
| **Systolic Blood Pressure** | Shared decision making application | Blood pressure will be reassessed at the end of Phase 2 | Blood Pressure | mm Hg | 60 | 200 | 1x baseline, 1x wk 11-12 | Good, provided it is measured correctly | | BloodPressure (systolic + diastolic) |
| **Diastolic Blood Pressure** | Shared decision making application | Blood pressure will be reassessed at the end of Phase 2 | Blood Pressure | mm Hg | 30 | 110 | 1x baseline, 1x wk 11-12 | Good, provided it is measured correctly | | BloodPressure |

### 3.2.2 Data Types of Clinical Characteristics

**Table 3 Parameters and measuring devices for category Clinical Characteristics**

| **Clinical Characteristics** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **HbA1C (glycated hemoglobin)** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | HbA1C | % | 2 | 15 | 1x baseline, 1x wk 11-13 | Depends on reliability of lab test | **Primary outcome** according to Power2DM Quantification Campaign Protocol (C.1 Power2DM Quantification Campaign Protocol LUMC v1_9 2016-07-21.pdf) | SimpleQuantity Observation |
| **Fasting glucose** | Clinical/Lab test, result uploaded to Shared decision making application | From lab, CGM | Fasting glucose | mmol/L | 3 | 25 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Fasting insulin** | Clinical/Lab test, result uploaded to Shared decision making application | From lab, estimated from KADIS | Fasting insulin | pmol/L | 15 | 250 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Insulin sensitivity (%HOMA-2 S)** | Clinical/Lab test, result uploaded to Shared decision making application | Based on fasting glucose/insulin | Insulin sensitivity (%HOMA-2 S) | % | 0 | 100 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Beta cell function (%HOMA-2-B)** | Clinical/Lab test, result uploaded to Shared decision making application | Based on fasting glucose/insulin | Beta cell function (%HOMA-2 B) | % | 0 | 100 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Inflammation (mg/l hs-CRP)** | Clinical/Lab test, result uploaded to Shared decision | From lab | Inflammation (mg/l hs-CRP) | mg/L | 0 | 4 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | making application | | | | | | | | | |
| **Non-esterified fatty acids** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | Non-estrefied fatty acids | mmol/L | 0 | 1 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Triglycerides** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | Triglycerides | mmol/L | 0 | 7 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Cholesterol** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | Cholesterol | mmol/L | 1 | 10 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **HDL Cholesterol** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | HDL Cholesterol | mmol/L | 0 | 3 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **LDL Cholesterol** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | LDL Cholesterol | mmol/L | 1 | 7 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Cholesterol Ratio** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | Cholesterol Ratio | | | | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Urine-Albumin** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | Urine-Albumin | mg/L | 0 | 300 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Creatinine** | Clinical/Lab test, result uploaded to Shared decision making application | From lab | Creatinine | micromol/ L | 60 | 200 | 1x baseline | Depends on reliability of lab test | | SimpleQuantity Observation |
| **Cortisol** | Clinical/Lab test, result uploaded to Shared decision making application | From lab, hair sample | Cortisol | pg/mg | 3 | 30 | 1x wk 11-12 | Depends on reliability of lab test | | SimpleQuantity Observation |

### 3.2.3 Data Types of Blood Glucose Level

**Table 4 Parameters and measuring devices for category Blood Glucose Level**

| Blood Glucose Level | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Blood Glucose Level** | Abbot FreeStyleLibre Flash Glucose Monitor | Abbot FreeStyleLibre FGM measures blood glucose level with sensor | Interstitial fluid glucose levels (FSL) | mmol/L | 1 | 20 | Continuously | Depends of reliability of FGM | **Primary outcome** according to Power2DM Quantification Campaign Protocol (C.1 Power2DM Quantification Campaign Protocol LUMC v1_9 2016-07-21.pdf) | TimeSeriesObservation |
| **Blood Glucose Level** | iHealth Wireless Smart Glucose-Monitoring system (iHSGM) | Additional blood samples need to be drawn by means of a finger-prick | Interstitial fluid glucose levels (FSL) | mmol/L | 1 | 20 | 8x a day for the first 72 hours, after that >= 2x a day | Depends of reliability of iHSGM | **Primary outcome** according to Power2DM Quantification Campaign Protocol (C.1 Power2DM Quantification Campaign Protocol LUMC v1_9 2016-07-21.pdf) | SimpleQuantity Observation |
| **Date and time of hyperglycemic episodes** | FGM, iHSGM | Times recorded when blood glucose level indicates hyperglycemic episode | Date time | dd-mm-yyyy hh:mm | | | Continuously | Good, easy to check correctness | | Observation (may need a specific observation profile) |
| **Frequency of hyperglycemic episodes** | FGM, iHSGM | Number of times recorded when blood glucose level indicates hyperglycemic episode | Frequency | Number of hyperglycemic episodes | | | Continuously | Good, easy to check correctness | | Observation (may need a specific observation profile) |
| **Magnitude of hyperglycemic episodes** | FGM, iHSGM | Blood glucose level when this indicates hyperglycemic episode | Interstitial fluid glucose levels (FSL) | mmol/L | 10 | 20 | Continuously | Depends on reliability of FGM/iHSGM | | Observation (may need a specific observation profile) |
| **Date and time of hypoglycemic episodes** | FGM, iHSGM | Times recorded when blood glucose level indicates hypoglycemic episode | Date time | dd-mm-yyyy hh:mm | | | Continuously | Good, easy to check correctness | | Observation (may need a specific observation profile) |
| **Frequency of hypoglycemic episodes** | FGM, iHSGM | Number of times recorded when blood glucose level indicates hypoglycemic episode | Frequency | Number of hypoglycemic episodes | | | Continuously | Good, easy to check correctness | | Observation (may need a specific observation profile) |
| **Magnitude of hypoglycemic episodes** | FGM, iHSGM | Blood glucose level when this indicates hypoglycemic episode | Interstitial fluid glucose levels (FSL) | mmol/L | 1 | 4 | Continuously | Depends on reliability of FGM/iHSGM | | Observation (may need a specific observation profile) |

### 3.2.4 Data Types of Eating Behaviour

**Table 5 Parameters and measuring devices for category Eating Behaviour**

| Eating Behaviour | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Unixtime** | POWER2DM mobile application | Determined at time of recording food/drink consumption by patients in POWER2DM mobile app | Unixtime | Seconds (10 digits) | | | Daily | Good, if correctly programmed | | DietaryIntakeLog |
| **Date** | POWER2DM mobile application | Determined at time of recording food/drink consumption by patients in POWER2DM mobile app | Date | yyyy-mm-dd | | | Daily | Good, if correctly programmed | | DietaryIntakeLog |
| **Time** | POWER2DM mobile application | Determined at time of recording food/drink consumption by patients in POWER2DM mobile app | Time | hhmm (hour, minutes) | | | Daily | Good, if correctly programmed | | DietaryIntakeLog |
| **Amount of carbohydrates** | POWER2DM mobile application: manual input | Carbohydrate consumption recorded by patients in mobile app | Carbohydrates | grams | 0 | 1000 | Continuously | Depends on whether it is correctly filled in by patient(s) | | DietaryIntakeLog |
| **Type of food/drink** | Not decided yet | | | | | | | | Not decided yet | DietaryIntakeLog |
| **Amount of bread units** | POWER2DM mobile application: manual input | Bread units logged by patients | Bread units (carbohydrates) | Bread units (grams) | | | Continuously | Depends on whether it is correctly filled in by patient(s) | | DietaryIntakeLog |
| **Amount of calories** | POWER2DM mobile application: manual input | Calorie consumption recorded by patients | Calories | kcal | 0 | 5000 | Continuously | Depends on whether it is correctly filled in by patient(s) | | DietaryIntakeLog |
| **Amount of calories** | FatSecret | FatSecret calorie counter | Calories | kcal | 0 | 5000 | Continuously | Depends on whether it is correctly filled in by patient(s) | | DietaryIntakeLog |
| **Quality/Nutritional value** | Not decided yet | | | | | | Continuously | Depends on whether it is correctly filled in by patient(s) | Not decided yet | DietaryIntakeLog |
| **Time of consumption** | POWER2DM mobile application: manual input | Time of registering food consumption by patients | Time | hh:mm | | | Continuously | Depends on whether it is correctly filled in by patient(s) | | DietaryIntakeLog |
| **Water consumption** | POWER2DM mobile application: manual input | Water consumption registered by patients | Water | litres | 0 | 5 | Continuously | Depends on whether it is correctly filled in by patient(s) | | DietaryIntakeLog |

### 3.2.5 Data Types of Exercise and Energy Expenditure

**Table 6 Parameters and measuring devices for category Energy Expenditure**

| Exercise and Energy Expenditure | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Average resting heart rate** | Fitbit HR Charge 2 | Heart rate tracker in Fitbit | Heart rate | Number of beats per minute | 60 | 100* | Continuously | ***Depends on reliability of the Fitbit HR Charge 2 and on whether the patient wears the wristband continuously | *Maximum for heart rate when at rest | TrackerHeartRateSummary |
| **Current heart rate** | Fitbit HR Charge 2 | Heart rate tracker in Fitbit | Heart rate | Number of beats per minute | 60 | 220** | Continuously | Same as above (***) | **Maximum for heart rate during exercise | TrackerHeartRateSummary |
| **Steps** | Fitbit HR Charge 2 | Accelerometer in Fitbit | Steps | Number of steps taken | | | Continuously | Same as above (***) | | ActivityTrackerSummary |
| **Distance** | Fitbit HR Charge 2 | Steps * average step-length | Distance | km | 0 | 50 | Continuously | Same as above (***) | | ActivityTrackerSummary |
| **Calories burned** | Fitbit HR Charge 2 | Calculated based on physical data, basal metabolic rate and exercise | Calories | kcal | 0 | 5000 | Continuously | Same as above (***) | | ActivityTrackerSummary |
| **Floors/Stairs climbed** | Fitbit HR Charge 2 | Altimeter in Fitbit | Elevation | Increments of 3,05m (10 ft) | 0 | 50 | Continuously | Same as above (***) | | ActivityTrackerSummary |
| **Duration** | Fitbit HR Charge 2 | Active time based on accelerometer measurement | Time active (duration) [hours, minutes] | hh:mm | 0 | 10 | Continuously | Same as above (***) | | ActivityTrackerSummary |

### 3.2.6 Data Types of Sleep

**Table 7 Parameters and measuring devices for category Sleep**

| Sleep | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Duration** | Fitbit HR Charge 2 | | Time (duration) [hours, minutes] | hh:mm | 0 | 24:00 | Continuously | Depends on reliability of the Fitbit HR Charge 2 and on whether the patient wears the wristband continuously | | SleepQualitySleepSummary |
| **Frequency** | iHealth Wave | iHealth "Edge" and presumably the "Wave" measures number of times taken a nap, slept, woken up | Sleep frequency | Number of times slept | | | Continuously | Depends on reliability of the iHealth Wave | Source of information: https://ihealthlabs.com/fitness-devices/ihealth-edge | SleepQualitySleepSummary |
| **Efficiency** | iHealth Wave | iHealth "Edge" and the "Wave" measures "sleep efficiency" | Sleep efficiency | % | 0 | 100 | Continuously | Depends on reliability of the iHealth Wave | | SleepQualitySleepSummary |
| **Quality/Disturbances** | Fitbit HR Charge 2 | Time restless/awake in between sleep time | Time (duration) [minutes] | mm | 0 | 60 | Continuously | Depends on reliability of the Fitbit HR Charge 2 and on whether the patient wears the wristband continuously | | SleepQualitySleepSummary |
| **Quality/Disturbances** | Not yet decided | A one item Visual Analogue Scale (VAS) (scale from 0 to 10) will be filled in daily during the monitoring periods | Sleep quality | Number on scale of 0 to 10 | 0 | 10 | Daily | Good, provided it is correctly filled in by patient(s) | | SleepQualitySleepSummary |

### 3.2.7 Data Types of Stress

**Table 8 Parameters and measuring devices for category Stress**

| Stress | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Unixtime** | POWER2DM mobile application | Determined at time of recording stress level by patients in mobile app | Unixtime | Seconds (10 digits) | | | Daily | Good, if correctly programmed | | MindState |
| **Date** | POWER2DM mobile application | Determined at time of recording stress level by patients in mobile app | Date | yyyy-mm-dd | | | Daily | Good, if correctly programmed | | MindState |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Time** | POWER2DM mobile application | Determined at time of recording stress level by patients in mobile app | Time | hhmm (hour, minutes) | | | Daily | Good, if correctly programmed | | MindState |
| **Stress level** | POWER2DM mobile application: manual input | A visual analogue color scale will be integrated into the app for completion with each insulin measure input. | Level of experiencing stress | Number on (colour) scale | | | 6x a Day | Depends on patient (reliability, correctness) | | MindState |

### 3.2.8 Data Types of Mood

**Table 9 Parameters and measuring devices for category Mood**

| **Mood** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Unixtime** | POWER2DM mobile application | Determined at time of recording mood by patients in POWER2DM app | Unixtime | Seconds (10 digits) | | | Daily | Good, if correctly programmed | | MindState |
| **Date** | POWER2DM mobile application | Determined at time of recording mood by patients in POWER2DM app | Date | yyyy-mm-dd | | | Daily | Good, if correctly programmed | | MindState |
| **Time** | POWER2DM mobile application | Determined at time of recording mood by patients in POWER2DM app | Time | hhmm (hour, minutes) | | | Daily | Good, if correctly programmed | | MindState |
| **Mood level** | POWER2DM mobile application: manual input | A cartoon based pictorial will be used to assess affect. In this pictorial, different mood states are displayed on two axis with the y axis representing the valence and the x axis representing the arousal of the mood. | Valence and arousal | Numbers on scales/axes | | | 6x a Day | Good, provided it is correctly filled in by patient(s) | | MindState |

### 3.2.9 Data Types of Diabetes Medical Treatment

**Table 10 Parameters and measuring devices for category Medical Treatment**

| **Diabetes Medical Treatment** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **Unixtime** | POWER2DM mobile application | Determined at time of recording medicine taken by patients in POWER2DM mobile app | Unixtime | Seconds (10 digits) | | | Daily | Good, if correctly programmed | | MedicationAdministration |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Date** | POWER2DM mobile application | Determined at time of recording medicine taken by patients in POWER2DM mobile app | Date | yyyy-mm-dd | | | Daily | Good, if correctly programmed | | MedicationAdministration |
| **Time** | POWER2DM mobile application | Determined at time of recording medicine taken by patients in POWER2DM mobile app | Time | hhmm (hour, minutes) | | | Daily | Good, if correctly programmed | | MedicationAdministration |
| **Type (oral/insulin)** | POWER2DM mobile application: manual input | Type of medication taken recorded by patients in POWER2DM mobile app | Type of medicine | "Oral" or "Insulin" | | | Daily | Good, provided it is correctly filled in by patient(s) | | MedicationAdministration |
| **Name of medicine** | POWER2DM mobile application: manual input | Name of medication taken recorded by patients in POWER2DM mobile app | Name of medicine | Text string | | | Daily | Good, provided it is correctly filled in by patient(s) | | MedicationAdministration |
| **Dosage/Amount** | POWER2DM mobile application: manual input | Dosage of medication taken recorded by patients in POWER2DM mobile app | Dosage | mg | 0 | 50 | Daily | Good, provided it is correctly filled in by patient(s) | | MedicationAdministration |

### 3.2.10 Data Types of Questionnaires

As discussed during POWER2DM consortium meeting in Salzburg on 8 and 9 February 2017, the questionnaires for prototype 1 will be selected by LUMC.

**Table 11 Parameters and measuring devices for category Questionnaires**

| **Questionnaires** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Parameter*** | ***Measuring device*** | ***More specific measuring device/method*** | ***Physical property to measure/ register*** | ***Units*** | ***Valid range lower cut-off*** | ***Valid range higher cut-off*** | ***How often and when data is measured/collected*** | ***(Expected) data quality*** | ***Comments*** | ***Mapping data type PDS model*** |
| **WHO-5 Well-Being Index (WHO)** | Questionnaire | WHO-5 Well-Being Index (WHO) | Well-Being Index | Number of points rated on (Likert) scales in questionnaire | 0 | 100 | 1x baseline, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | Five items are assessed on a 6-point Likert scale ranging from 0 to 5 and the individual item scores are added together and transformed into a 100 point scale. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Patient Health Questionnaire (PHQ-9)** | Questionnaire | Patient Health Questionnaire (PHQ-9) | Patient Health | Number of points rated on (Likert) scales in questionnaire | 0 | 27 | 1x baseline, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | 9 DSM-IV criteria for depression on a four point (0-3) scale. The sum of these scores (range= 0-27) is then used to assess the level of depression | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Brief measure of Generalized Anxiety Disorder (GAD-7)** | Questionnaire | Brief measure of Generalized Anxiety Disorder (GAD-7) | Measure of Generalized Anxiety Disorder | Number of points rated on (Likert) scales in questionnaire | 0 | 21 | 1x baseline, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | A seven item anxiety scale (0-3) that measures how often the patient has been bothered by specific symptoms of anxiety over the previous two weeks. The sum of these scores (range= 0-21) is then used as an assessment of the presence of anxiety | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Perceived Stress Scale (PSS)** | Questionnaire | Perceived Stress Scale (PSS) | Perceived Stress | Number of points rated on (Likert) scales in questionnaire | 0 | 40 | 1x baseline, 1x wk 4, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | 10 item scale designed to measure the degree to which situations' in one's life are appraised as stressful on a 5 point scale. All scores are summed for an overall perceived stress score | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Problem Areas in Diabetes (PAID)** | Questionnaire | Problem Areas in Diabetes (PAID) | Problem Areas in Diabetes | Number of points rated on (Likert) scales in questionnaire | 0 | 100 | 1x baseline, 1x wk 4, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | 20-item measure of psychosocial adjustment to diabetes rated on a 5 point Likert scale (range = 0-4). Added together and multiplied by 1.25 to transform the raw score into a 0-100 scale with higher scores representing increased emotional distress | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Diabetes Self-Management Questionnaire (DSMQ-R)** | Questionnaire | Diabetes Self-Management Questionnaire (DSMQ-R) | Diabetes Self-Management | Number of points rated on (Likert) scales in questionnaire | | | 1x baseline, 1x wk 4, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | 20 items about self-management. These items can be grouped into four factors (glucose management, dietary control, physical activity, and physician contact) from which a sum score can be calculated for overall glycemic control. All subscales are highly correlated with levels of HbA1c. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Audit of Diabetes Dependent QoL (ADDQoL)** | Questionnaire | Audit of Diabetes Dependent QoL (ADDQoL) | Diabetes Dependent Quality of Life | Number of points rated on (Likert) scales in questionnaire | | | 1x baseline, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | The ADDQoL was developed to measure individuals' perceptions of the impact of diabetes on their quality of life in 13 domains on a 6 point scale for how much better this domain of their life would be if they didn't have diabetes and how important that aspect of their life is on a 4 point scale. Additionally, there are two items assessing perceived general quality of life and imagined general quality of life without diabetes. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Hypoglycemia Fear Survey (HFS)**** | Questionnaire | Hypoglycemia Fear Survey (HFS) | Fear of Hypoglycemia | Number of points rated on (Likert) scales in questionnaire | | | 1x baseline | Good, provided it is correctly filled in by patient(s) | 37 item survey designed to quantify the fear associated with hypoglycemia in diabetes patients. The survey is divided into two sections assessing specific behaviours used to avoid low blood sugar and worry associated with low blood sugar.<br>** Only participants who use insulin, medication associated with hypoglycemic episodes, or those indicated as potentially having hypoglycemic related distress based on their responses on the PAID will be required to complete this questionnaire | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Diabetes Eating Problem Survey (DEPS-R)*** | Questionnaire | Diabetes Eating Problem Survey (DEPS-R) | Diabetes Eating Problem | Number of points rated on (Likert) scales in questionnaire | 0 | 80 | 1x baseline | Good, provided it is correctly filled in by patient(s) | 16 items (6 point Likert scale (0-5)) to assess the presence of eating related problems in T1DM patients.<br>* This measure will only be used if a patient engages in an associated self-management task (e.g. only insulin users will be asked about anxiety related to using insulin) or they indicate associated problems in other questionnaires (e.g. DEPS-R will be administered if the patient indicates issues regarding eating) | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Fear of Complications Questionnaire (FCQ)*** | Questionnaire | Fear of Complications Questionnaire (FCQ) | Fear of Complications | Number of points rated on (Likert) scales in questionnaire | 0 | 45 | 1x baseline | Good, provided it is correctly filled in by patient(s) | 15 item questionnaire that evaluates patient fear surrounding long-term complications of diabetes on a four point scale. A sum score of all the items can be used to calculate overall fear of complications.<br>* This measure will only be used if a patient engages in an associated self-management task (e.g. only insulin users will be asked about anxiety related to using insulin) or they indicate associated problems in other questionnaires | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Diabetes Fear of Injecting and Self-Testing Questionnaire (D-FISQ)***** | Questionnaire | Diabetes Fear of Injecting and Self-Testing Questionnaire (D-FISQ) | Diabetes Fear of Injecting and Self-Testing | Number of points rated on (Likert) scales in questionnaire | | | 1x baseline | Good, provided it is correctly filled in by patient(s) | 21 item questionnaire designed to assess fear of injecting insulin and fear of self-testing blood glucose in diabetes patients [62]. The patient reports on the frequency of given actions on a 4 point scale<br>*** The questionnaire can be split into specific scales to measure the two individual factors of fear of injecting insulin and fear of self-testing blood glucose. Patients who indicate on the PAID that they may have anxiety related to self-testing will be required to complete the fear of self-testing blood glucose scale, and only those on insulin therapy or who indicate on the PAID anxiety related to insulin injections will be required to fill in the fear of injecting insulin scale. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **After-Scenario Questionnaire (ASQ)** | Questionnaire | After-Scenario Questionnaire (ASQ) | After-Scenario Questionnaire (ASQ) | Number of points rated on (Likert) scales in questionnaire | | | 1x wk 1, 1x wk 4 | Good, provided it is correctly filled in by patient(s) | The user rates their agreement with three satisfaction statements assessing ease, time, and support of use on a 7 point scale ranging from strongly agree (1) to strongly disagree (7) with the option of choosing not applicable. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Clarke score** | Questionnaire | Clarke score | hypo-(un)awareness | 8 multiple choice questions | | | 1x baseline, 1x wk 4, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | Only for T1DM patient. Each one of these (Clarke score and Gold score) questionnaires has specific ways to calculate the cut-off scores to indicate the presence or absence of problems. Ideally, we would automate the calculations so only the indication is shown. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |
| **Gold score** | Questionnaire | Gold score. Questions patient: “Do you know when your hypos are commencing?” | hypo-(un)awareness | 7-point Likert scale, with 1 representing “always aware” and 7 representing “never aware”. | 1 | 7 | 1x baseline, 1x wk 4, 1x wk 11-12 | Good, provided it is correctly filled in by patient(s) | Only for T1DM patient. A score of ≥4 implies impaired awareness of hypoglycemia. | QuestionnaireResponse (holds the actual answers), QuestionnaireResult (the score calculated from the answers) |

# 4 DATA PROCESSING

Based on the data quality analysis framework as described in chapter 3, the data quality checks have been implemented into the data integration platform. The data is checked before it is saved to the Personal Data Store (PDS). The business rules including the lower and upper cut off values for the data types used have been implemented. The development process has been as follows:

1. Definition of all business rules and associated parameters in configuration;
2. Implementation of data validation based on business rules and parameters defined;
3. Integration with PDS according to interface specification;
4. Unit testing of the data validation class.

The following subsections give a detailed overview of the results of each of the above steps.

## 4.1 Configuration

The configuration datafile that contains all the parameters and business rules criteria to validate incoming data:

```
{     "Parameters": [
        {
          "group": "Patient Characteristics",
          "description": "",
        "parametername": "age",
        "low": "10",
        "high": "111",
        "unit": "years"
      },

      {
        "group": "Patient Characteristics",
        "description": "",
        "parametername": "weight",
        "low": "10",
        "high": "250",
        "unit": "kg"
      },

      {
        "group": "Patient Characteristics",
        "description": "",
        "parametername": "height",
        "low": "100",
        "high": "220",
        "unit": "cm"
      },

      {
        "group": "Patient Characteristics",
        "description": "male=0 female=1",
        "parametername": "gender",
        "low": "0",
        "high": "1",
        "unit": ""
      },

      {
        "group": "Patient Characteristics",
```

```
  "description": "T1DM=1 T2DM=2",
  "parametername": "typediabetes",
  "low": "1",
  "high": "2",
  "unit": ""
},

{
  "group": "Patient Characteristics",
  "description": "",
  "parametername": "waist",
  "low": "40",
  "high": "175",
  "unit": "cm"
},

{
  "group": "Patient Characteristics",
  "description": "",
  "parametername": "waisthipratio",
  "low": "0.6",
  "high": "1.4",
  "unit": ""
},

{
  "group": "Patient Characteristics",
  "description": "",
  "parametername": "bmi",
  "low": "10",
  "high": "75",
  "unit": "kg/m2"
},

{
  "group": "Patient Characteristics",
  "description": "",
  "parametername": "systolicbloodpressure",
  "low": "60",
  "high": "200",
  "unit": "mm HG"
},

{
  "group": "Patient Characteristics",
  "description": "",
  "parametername": "diastolicbloodpressure",
  "low": "30",
  "high": "110",
  "unit": "mm HG"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "hba1c",
  "low": "2",
  "high": "15",
  "unit": "%"
},
```

```
{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "fastingglucose",
  "low": "3",
  "high": "25",
  "unit": "mmol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "fastinginsulin",
  "low": "15",
  "high": "20",
  "unit": "pmol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "insulinsensitivity",
  "low": "0",
  "high": "100",
  "unit": "%"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "betacellfunction",
  "low": "0",
  "high": "100",
  "unit": "%"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "inflamation",
  "low": "0",
  "high": "4",
  "unit": "mg/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "nonestrefieddfattyacids",
  "low": "0",
  "high": "1",
  "unit": "mmol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "tryglycerides",
  "low": "0",
  "high": "7",
  "unit": "mmol/L"
```

```
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "cholesterol",
  "low": "1",
  "high": "10",
  "unit": "mmol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "hdlcholesterol",
  "low": "0",
  "high": "3",
  "unit": "mmol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "ldlcholesterol",
  "low": "1",
  "high": "7",
  "unit": "mmol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "urinealbumin",
  "low": "0",
  "high": "300",
  "unit": "mg/L"
},


{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "creatininine",
  "low": "60",
  "high": "200",
  "unit": "micromol/L"
},

{
  "group": "Clinical Characteristics",
  "description": "",
  "parametername": "cortisol",
  "low": "3",
  "high": "30",
  "unit": "pg/mg"
},

{
  "group": "Blood Glucose Level",
  "description": "",
  "parametername": "interstitialfluidglucoselevels_ihealth",
```

```
        "low": "1",
        "high": "20",
        "unit": "mmol/L"
      },

      {
        "group": "Blood Glucose Level",
        "description": "",
        "parametername": "interstitialfluidglucoselevels_fsl",
        "low": "1",
        "high": "4",
        "unit": "mmol/L"
      },

      {
        "group": "Blood Glucose Level",
        "description": "",
        "parametername":
"interstitialfluidglucoselevels_hyperglycemicepisode",
        "low": "10",
        "high": "20",
        "unit": "mmol/L"
      },

      {
        "group": "Blood Glucose Level",
        "description": "",
        "parametername":
"interstitialfluidglucoselevels_hypoglycemicepisode",
        "low": "1",
        "high": "4",
        "unit": "mmol/L"
      },

      {
        "group": "Eating behavior",
        "description": "",
        "parametername": "carbohydrates",
        "low": "0",
        "high": "1000",
        "unit": "grams"
      },

      {
        "group": "Eating behavior",
        "description": "",
        "parametername": "calories_recordedbypatient",
        "low": "0",
        "high": "5000",
        "unit": "kcal"
      },

                {
        "group": "Eating behavior",
        "description": "",
        "parametername": "calories_fastsecretcaloriecounter",
        "low": "0",
        "high": "5000",
        "unit": "kcal"
      },
```

```
{
  "group": "Eating behavior",
  "description": "",
  "parametername": "water",
  "low": "0",
  "high": "5",
  "unit": "litres"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "heartrate_rest_fitbit",
  "low": "60",
  "high": "100",
  "unit": "beats/minute"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "heartrate_exercise_fitbit",
  "low": "60",
  "high": "220",
  "unit": "beats/minute"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "distance",
  "low": "0",
  "high": "50",
  "unit": "km"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "calories",
  "low": "0",
  "high": "5000",
  "unit": "km"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "elevation",
  "low": "0",
  "high": "50",
  "unit": "3.05m"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "timeactive",
  "low": "0",
  "high": "10",
  "unit": "khh:mm"
```

```
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "calories_ihealth",
  "low": "0",
  "high": "5000",
  "unit": "kcal"
},

{
  "group": "Exercise and energy expenditure",
  "description": "",
  "parametername": "distance_ihealth",
  "low": "0",
  "high": "10",
  "unit": "km"
},

{
  "group": "Sleep",
  "parametername": "sleeptime",
  "description": "",
  "low": "0",
  "high": "24",
  "unit": "hh:mm"
},

{
  "group": "Sleep",
  "parametername": "sleepefficiency",
  "description": "",
  "low": "0",
  "high": "100",
  "unit": "%"
},

{
  "group": "Sleep",
  "description": "",
  "parametername": "awaketime",
  "low": "0",
  "high": "60",
  "unit": "mm"
},

{
  "group": "Sleep",
  "description": "",
  "parametername": "sleepquality",
  "low": "0",
  "high": "10",
  "unit": ""
},

{
  "group": "Stress",
  "description": "",
  "parametername": "respiratoryrate",
  "low": "10",
```

```
    "high": "50",
    "unit": "breath/minute"
  },

  {
    "group": "Mood",
    "description": "",
    "parametername": "noentry",
    "low": "0",
    "high": "0",
    "unit": ""
  },

  {
    "group": "Medical treatment",
    "description": "",
    "parametername": "dosage",
    "low": "0",
    "high": "50",
    "unit": "mg"
  },

  {
    "group": "Questionaires",
    "description": "",
    "parametername": "wellbeingindex",
    "low": "0",
    "high": "100",
    "unit": ""
  },

  {
    "group": "Questionaires",
    "description": "",
    "parametername": "patienthealth",
    "low": "0",
    "high": "27",
    "unit": ""
  },


  {
    "group": "Questionaires",
    "description": "",
    "parametername": "anxietydisorder",
    "low": "0",
    "high": "21",
    "unit": ""
  },

  {
    "group": "Questionaires",
    "description": "",
    "parametername": "perceivedstress",
    "low": "0",
    "high": "40",
    "unit": ""
  },

  {
    "group": "Questionaires",
```

```
        "description": "",
        "parametername": "diabetesproblemareas",
        "low": "0",
        "high": "100",
        "unit": ""
      },

      {
        "group": "Questionaires",
        "description": "",
        "parametername": "diabeteseatingproblem",
        "low": "0",
        "high": "80",
        "unit": ""
      },

      {
        "group": "Questionaires",
        "description": "",
        "parametername": "fearofcomplications",
        "low": "0",
        "high": "45",
        "unit": ""
      },

      {
        "group": "Questionaires",
        "description": "",
        "parametername": "hypoawareness",
        "low": "1",
        "high": "7",
        "unit": ""
      }

      ]
}
```

## 4.2 Data validation implementation

The actual code that validates incoming data based on the configuration as presented in the previous section:

```
package datavalidation.impl;

import java.io.FileReader;

import com.google.gson.JsonArray;
import com.google.gson.JsonElement;
import com.google.gson.JsonObject;
import com.google.gson.JsonParser;

import datavalidation.DataValidationResult;
import datavalidation.IP2DMDatavalidation;

public class P2DMDatavalidation implements IP2DMDatavalidation {

      private JsonParser parser;
      private JsonObject jsonobj;
      private JsonArray jsonArray;
```

```
    P2DMDatavalidation() {

      try
      {
        parser = new JsonParser();


        jsonobj        =        parser.parse(        new        FileReader(
"D:\\Power2DM\\Repositories\\horizon-2020_power2dm.datavalidation-
service\\P2DMDatavalidation\\src\\main\\resources\\Parameters.json"       )
).getAsJsonObject();
        jsonArray = jsonobj.getAsJsonArray("Parameters");
      }
      catch  (Exception e)
        {
            System.out.println(   "Error   creating   datavalidation:   "   +
e.getMessage() );
        }
    }

      public         DataValidationResult         validateObservation(String
observationCode, String observationUnit,
                  Double observation) {

            DataValidationResult           result           =           new
DataValidationResult("Observationcode not found...");

            try {

            // Look for observationcodes in JSON...
                  for ( JsonElement element : jsonArray ) {

                        JsonObject jsonObject = element.getAsJsonObject();

                        String     parameterName     =     jsonObject.get(
"parametername" ).getAsString().toUpperCase();

                        // Check if observation is in range...
                        if              (              parameterName.equals(
observationCode.toUpperCase() ) ) {

                              System.out.println(    "Parametername:    "    +
parameterName );

                                double    low    =    jsonObject.get(    "low"
).getAsDouble();
                                System.out.println( "low: " + low );

                                double    high    =    jsonObject.get(    "high"
).getAsDouble();
                                System.out.println( "high: " + high );

                                if ( ( observation >= low ) && ( observation
<= high ) )  {
                                      result.setValid( true );
                                      System.out.println(  "Observation  "  +
parameterName + " with value " + observation +", is valid");
                                }
                                  else {
```

```
                                    result.setValid(false);//          new
DataValidationResult("Value of observation " + observationCode + "  out of
bounds. ");  // + ' Value: ' + observation );
                                    result.setErrorMessage("Value       of
observation " + observationCode + "  out of bounds.");
                                    System.out.println( "Observation  " +
parameterName + " with value " + observation +", is NOT valid");
                                     }

                              System.out.println( "---");

                              break;
                        }

                        // Check unit...
                        String     unit     =     jsonObject.get(     "unit"
).getAsString();
                        System.out.println( "Unit: " + unit );

                        // check on the unit
//                      if    (    !observationUnit.equals(    unit    )    &&
result.isValid() ) {
//                            result.setValid( false );
//                            result.setErrorMessage("Unit        "        +
observationUnit + "differs from unit in database: " + unit );
//                      }
                  }

                  // Give back result....

                  return result;

                  } catch (Exception e) {
                        System.out.println( "Error: " + e.getMessage() );
                  }

        return result;
      }
}
```

## 4.3 Integration with the Personal Data Store

Interface specification that the data validation class must implement to establish integration with the PDS:

```
package datavalidation;

public interface IP2DMDatavalidation {

    /**
     * Validate a measurement if it is in normal ranges
     * @param observationCode Valueset for observations are defined in
http://app.srdc.com.tr/power2dm/pds/ValueSet/pdm-observations
     *                       (Which of these observations will be covered
in Data Validation ?? )
     * @param observationUnit Unit of observation (In the PDS data model,
for each observation type, a unit is specified)
     * @param observation The value of observation
     * @return Validation result; a boolean indicating the result and error
message if validation is failed
```

```
     */
    public DataValidationResult validateObservation(String observationCode,
String observationUnit, Double observation);

}
```

## 4.4 Unit test

Unit test for the data validation class implementation:

```
package datavalidation;

public class DataValidationResult {
    private boolean valid;
    private String errorMessage;

    public DataValidationResult(){
        valid = true;
        errorMessage = "";
    }

    public DataValidationResult(String errorMsg){
        valid = false;
        errorMessage = errorMsg;
    }

    public boolean isValid() {
        return valid;
    }

    public void setValid(boolean valid) {
        this.valid = valid;
    }

    public String getErrorMessage() {
        return errorMessage;
    }

    public void setErrorMessage(String errorMessage) {
        this.errorMessage = errorMessage;
    }

}
```