

Standardizing Codes in PERSIMUNE Lab Results from MedCom & RegionH

Background

PERSIMUNE is currently collecting lab results from about 35,000 patients. From the biochemistry and microbiology result we have extracted 5,337 different combinations of analysis codes and codes identifying the responsible lab (Excel file with all analysis code data attached as 5.6_PERSIMUNE_Lab-codes.xlsx).

Analysis codes can be NPU codes, but many are local codes and they are not unique across labs, so no automated sorting is possible without normalization.

An analysis to find e.g. evidence for a specific infection can be represented by a number of NPU codes and local codes. Some examples: In a sample of 350 patients with 1100 analysis HCV-RNA is represented with 9 different analysis codes and 13 some immune globulin analyses are represented by 50 different codes.

Standardizing data to be analyzable

We need normalizing in several different levels:

Standards

We need to decide on naming convention for analysis, ID-codes of analysis, labs, materials, and units

Positive/negative values

For some analysis we only require results indicated as a positive/negative value. It is possible to extract this information from 5-6 different fields in the MedCom data to determine if the result is positive or negative. This is not a trivial task because the result can be hidden in a longer comment or has to be calculated to be between an upper and lower limit or above/below a certain threshold. Each analysis needs to be treated differently and experts have to define algorithms.

Quantitative values

Identification of analysis methods to conclude on comparability of results from different labs is important. Likewise units and material differs a lot and needs to be normalized or marked uniformly. The same unit/material can be written in different ways and typos don't make it easier. This calls for decision on a standard that can be implemented in the analysis dataset.

Tools and process

We suggest establishing a coordinating group to

- divide the list of analysis in disease area lists
- ensure uniformity of normalizing across disease areas by suggesting some standards e.g. for coding of labs, material and units and provide guidance on coding for the disease area groups
- keep oversight with the progress

Suggested members: Marie Helleberg, Ruth Frikke-Schmidt, Lars Peters, Søren Schwartz Sørensen, Erik Viuff Hansen

A number of disease area specialist groups to be responsible for normalizing coding within a subset of analysis, e.g.

- Hematology/oncology (incl. neutropenia/leucopenia), Henrik Sengeløv & Gedske Daugaard to suggest team

- Nephrology, Søren Schwartz Sørensen and Marianne Rix to suggest team
- Hepatology, Allan Rasmussen, Henriette Ytting Lambert and Lars Peters to suggest team
- Infections, including HIV & CF, Jan Gerstoft and Therese Katzenstein to suggest team
- Rheumatology, Søren Jacobsen to suggest team
- Clinical immunology, Sisse Rye Ostrowski to suggest team

To perform the standardizing process the PERSIMUNE has developed a tool to work within the PERSIMUNE web interface.

Typelister

Typelister

Vælg liste: Tabel analyse units

Opdater listen

ID	RegUnit	DisplayUnit	☐ Er kontrolleret (skjul markerede felter)
96	U/l	U/L	☐
97	µg/mL	µg/mL	☐
98	10E3 IU/L	10 ³ IU/L	☐
99	10E3 kp/L	10 ³ kp/L	☐
100	genotype	genotype	☐
101	IU/L	IU/L	☐
102	IU/ml	IU/mL	☐
103	%	%	☐
104	g/l	g/L	☐

The tool is supposed to be flexible. It can be set up to group or standardize different parameters. The process flow will be:

1. Decide which parameters to modify, and who should be allowed to do it
2. IT sets up the table and grant access to the parameters to be modified.
3. The authorized person(s) write the changes into the webpage and save them.
4. Different scripts implement the changes to the relevant tables.

Erik Viuff Hansen
Jesper Grarup
08 September 2015