EDUPS Formulary (NL) – Technical description

# Introduction

The Edups Formulary (NL) Database is a relational database designed to manage and organize drug information within Edups. This database serves as a centralized repository for storing details about various drugs, their active ingredients, dosage forms, strengths, and associated brands.

The code and files referred to in this document are found in the [repository](https://github.com/FarmaSync/edups).

# Data Origin

## Data Sources

The data integrated into the Formulary Database originates from the database provided by the ‘College ter Beoordeling van Geneesmiddelen’ (CBG), the Dutch regulatory body for drugs (equivalent to the EMA/FDA).

This database contains information for all drugs authorized in the Netherlands, including the trade name, dosage forms and active ingredients.

# Data Preparation Steps

1. **Filtering based on dispensing status**

Only drugs that require a prescription were included in our dataset

1. **Strength extraction**

The strengths of the drugs were extracted from the tradename.[[1]](#footnote-1) A strength (or concentration) was extracted in 98% of the cases. In the remaining 2% no strength was present, or the description was too complex. **These drugs were not included in the dataset**. Below are some observations and points to be aware of:

|  |  |
| --- | --- |
| **Cat.** | **Description** |
| Multiple components (1) | In the case of drugs with multiple active components, multiple strengths (or concentrations) are present.[[2]](#footnote-2) Separating these does not provide any benefit at this stage. **Therefore, they are not separated, and the strength-amount and strength-unit are not separated either.** |
| Multiple components (2) | In the case of multiple active component drugs with the same strength-unit, nomenclature can vary, sometimes showing the unit once, and other times showing it for each component.[[3]](#footnote-3) Standardizing this is possible but will cost more time and does not add much benefit at this stage. Therefore, **these cases are not further standardized.** |

1. **Strength unit standardization**

A total of 61 distinct units were identified. Sometimes, these are expressed using different terms. For example, "microgram" can also be written as "mcg" or "µg." This was standardized according to a standardization table.[[4]](#footnote-4)

1. **Strength amount decimal separator standardization**

All decimal separators were standardized to a comma, while thousand separators were standardized to a point.

1. **Active ingredient cleaning and processing**

The active ingredients sometimes contain salt and hydration terms. These were identified and stripped out[[5]](#footnote-5). Below are some observations and points to be aware of:

|  |  |
| --- | --- |
| **Cat.** | **Description** |
| Duplicates | Drugs appear multiple times, sometimes alone and other times in combination with different substances (e.g., ‘Amoxicilline’ and ‘Amoxicilline/clavulaanzuur’) |
| Confusing components | Complex (combinations of) specific strains or ions (e.g., ‘A-Thailand-8-2022 (H3N2)’ and ‘(Ca²⁺)/(E 509)/(CL⁻)’) |

1. **Dosage form extraction**

The dosage forms were extracted from the tradename[[6]](#footnote-6). The dosage forms are specific and numerous (n = 153), but already normalized. **Most drugs are covered by a limited number of dosage forms** (75% of the drugs are covered by the top 15 dosage forms). For this reason, we adhere to the current nomenclature and no further standardization is applied.

# Database overview

The Formulary Database is available in the repository. A demo is accessible through: <https://edupsformulary.streamlit.app>

## Schema Diagram

Afbeelding met schermopname, tekst, Lettertype, zwart

Automatisch gegenereerde beschrijving

## Maintenance

Once implemented, the IDs used cannot be changed anymore. Therefore, any adjustments to the database (e.g., insertion or deletion of a drug) should be applied to the database itself and not through re-creation of the database.

To add a new drug, the user must ensure that the active ingredient and dosage form exist. If not, these need to be inserted first. **It is unlikely that any adjustments need to be made in the coming years.**

# Appendix

## Strength Unit standardization

Where applicable, strength units were standardized according to the table below.

|  |  |
| --- | --- |
| **Variant** | **Standard** |
| ppm | Ppm |
| µg | mcg |
| microgram |
| microgam |
| mcg |
| mg | mg |
| miligram |
| g | g |
| gram |
| microliter | microliter |
| ml | ml |
| mililiter |
| l | l |
| liter |
| IE | IE |
| I.E |
| internationale eenheden |
| IU |
| I.U |
| miljoen IE | miljoen IE |
| miljoen I.E. |
| miljoen internationale eenheden |
| eenheden | eenheden |
| E |
| U |
| miljoen E | miljoen eenheden |
| ME |
| MU |
| DU | DU |
| AU | AU |
| **Variant** | **Standard** |
| SQ-E | SQ eenheden |
| anti-Xa-eenheden | anti-Xa-eenheden |
| Allerganeenheden | Allerganeenheden |
| dosis |  |
| plaquevormende eenheden (PFU)/ml | PFU |
| PFU |
| plaquevormende eenheden |
| BU | BU |
| Speywood-eenheden | Speywood-eenheden |
| KIE | KIE |
| anti-heparine IE | anti-heparine IE |
| uur | Uur |
| % | % |
| mmol | mmol |
| mol | mol |
| Bq | Bq |
| MBq | MBq |
| kBq | kBq |
| SQ-T | SQ-T |
| -SQ-T |
| SQ-HDM | SQ-HDM |
| SQ-Bet | SQ-Bet |
| TE | TE |
| sferoïden | sferoïden |
| cm² | cm² |
| IR | IR |
| vg | vg |
| cellen | cellen |

1. E.g., ‘1 mg/ml’ from the tradename ‘“Monofree Dexamethason **1 mg/ml**, oogdruppels, oplossing in verpakking voor éénmalig gebruik”’. [↑](#footnote-ref-1)
2. E.g., ‘Sitagliptine/Metformine Teva 50 mg/1000 mg, filmomhulde tabletten’ [↑](#footnote-ref-2)
3. E.g., ‘50 mg/1000 mg vs. ‘150/37,5/200 mg’ [↑](#footnote-ref-3)
4. Refer to **Strength Unit standardization**. [↑](#footnote-ref-4)
5. E.g., ‘DEXAMETHASONDINATRIUMFOSFAAT’ becomes ‘DEXAMETHASON’. [↑](#footnote-ref-5)
6. E.g., ‘**oogdruppels’** from the tradename ‘“Monofree Dexamethason 1 mg/ml, **oogdruppels**, oplossing in verpakking voor éénmalig gebruik”’. [↑](#footnote-ref-6)