Regulatory Toxicology and Pharmacology 98 (2018) 9–17

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Development of nutritional risk assessment platform in Korea

Hunjoo Lee[a], Nam ji Kwon[a,b], Yongsoo Kim[c], Eunyoung Han[b,d,*]

[a] CHEM.I.NET Ltd., Room 302, 773-3, Mok-dong, Yangcheon-gu, Seoul, South Korea
[b] College of Pharmacy, Duksung Women's University, Seoul, South Korea
[c] Korea Health Industry Development Institute, 187 Osongsaengmyeong2(i)-ro, Gangoe-myeon, Cheongwon-gun, Chungcheongbuk-do, 363-951, South Korea
[d] Innovative Drug Center, Duksung Women's University, 33, Samyangro 144-gil, DobongGu, Seoul, South Korea

ARTICLE INFO

Keywords:
Nutritional risk assessment
Integrated database
Dietary modeling
Simulation
Nutritional fortification and reduction

ABSTRACT

Risk assessment has been used to prevent health problems associated with eating habits in response to increased interest in a balanced diet. For nutritional risk assessment (NRA), it is important to 1) consider personal nutrition status based on year-round dietary intake, 2) organize core datasets such as food composition, intake, and health based guidance value (HBGV) datasets with public confidence, and 3) assess and predict the effects by using the computerized NRA tool. Our research staff constructed an integrated database system by compiling and organizing core datasets produced sporadically by different organizations and with different formats and developed a nutritional exposure and risk assessment system called Nutri-Risk (NUTRItional RISK assessment platform), which contained the database. Nutri-Risk is not only capable of NRA, but also contains additional data service functions. Here, the compilations and organization of an integrated database are outlined. In addition, the overall architectures of Nutri-Risk and dietary modeling are described and predictive simulation functions to support the regulatory decisions related to nutritional fortification or reduction policy were demonstrated.

1. Introduction

Interest in a balanced diet has increased because of health problems associated with excess or deficit consumption of nutrients. Ahmed and Haboubi (2010) and Daboné et al. (2011) reported that, because people have health problems caused by both malnutrition and over-nutrition, it is essential to conduct nutritional risk assessment (NRA) to investigate the nutritional status of individuals and develop appropriate ways to treat their problems.

Nutritional risk assessment, which focuses on making decisions to protect health in the face of scientific uncertainty, can be generally described as “characterizing the potential hazards and associated risks to life and health resulting from exposure of humans to chemicals present in food over a specified period” (WHO, 2009). Some aspects of the risk assessment approach used for hazardous chemicals (e.g., contaminants, pesticides, etc.) and non-nutrients are applicable to the assessment of risks related to nutrition. The collection of core datasets such as food composition, intake, and health based guidance value (HBGV) datasets with public confidence for a target nutrient and the refined organization of different datasets are prerequisite to obtaining a more realistic estimation of exposure and risk assessment. It is also necessary to conduct a scientific and transparent assessment process that considers the population-level variability among diets. The exposure levels estimated by NRA may result in regulatory agents developing new nutritional policies such as fortification or reduction of nutrient amounts for foods or food groups. In this NRA process, the introduction of dietary modeling is needed to select effective foods or food groups (EFFG) that are expected to control exposure for all population groups.

Dietary modeling is a series of processes in which exposure values are calculated by combining specific food composition values with food consumption data, the contribution of exposure is evaluated, and EFFG that are expected to have considerable impacts on approval and release of new health functional foods (dietary supplement) and on increase or decrease of nutrient amounts in conventional foods need to be selected. It is important to conduct simulations before implementing regulations for tolerance levels of EFFG because the actual effects will be different from the expected effects (Fig. 1.).

However, there are compatibility issues because the core datasets have been produced sporadically by different organizations (i.e., governmental agencies and universities etc.) and in different formats in Korea. Moreover, development of automated and computerized tools for NRA modeling and simulation are required to handle the large amounts of information in the core datasets more efficiently.

To solve these issues and provide a useful software tool, our research staff built an integrated database by compiling and organizing

* Corresponding author. College of Pharmacy, Duksung Women's University, Seoul, South Korea.
E-mail addresses: adstar@cheminet.kr (H. Lee), njkwon@cheminet.kr (N.j. Kwon), kimys@khidi.or.kr (Y. Kim), homepage2600@duksung.ac.kr (E. Han).

https://doi.org/10.1016/j.yrtph.2018.06.022
Received 16 June 2017; Received in revised form 19 November 2017; Accepted 30 June 2018
Available online 02 July 2018

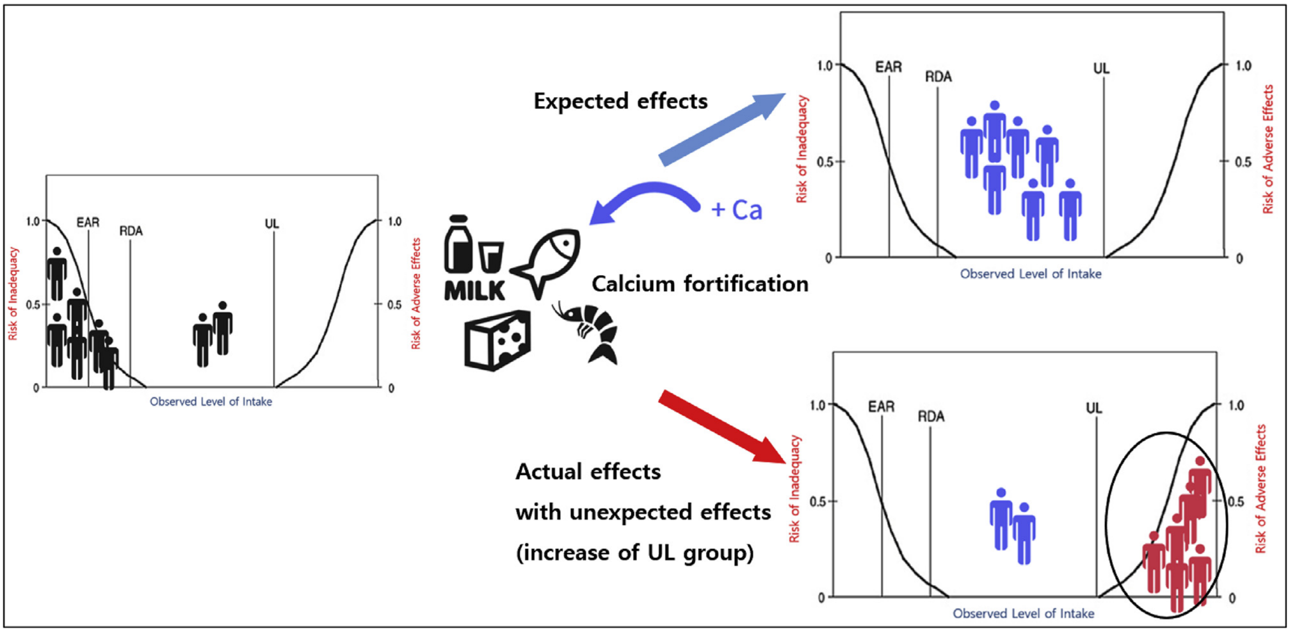

Fig. 1. The investigation of differences between actual effects and expected effects by the simulation.
* EAR: Estimated Average Requirements, RDA: Recommended Dietary Allowances, UL: Upper Limit.
* The simulation of nutritional fortification can expect the differences between actual effects and expected effects, and then the simulation supports efficient nutritional fortification.

the core datasets and developed a nutritional exposure and risk assessment system named Nutri-Risk (NUTRItional RISK assessment platform) embedding the database.

This study describes the compilation processes and structures of an integrated database as well as provides a detailed description of the architecture and overall features of Nutri-Risk, demonstrates the functions for systemic dietary modeling and predictive simulation implemented in Nutri-Risk, and discusses future directions of Nutri-Risk.

2. Method

2.1. Core datasets collection

To build an integrated database for core datasets, a total of 25 different datasets produced by Korean governmental authorities and universities from 2001 to 2016 were collected. The core datasets were categorized into the following three groups: food intake, food composition, and HBGV.

The food intake dataset consisted of a total of 17 (n = 2,531,253) studies, including epidemiological (cross-sectional and cohort) and intervention studies. There were a total of nine cross-sectional datasets (n = 2,210,069), including datasets produced from the Korea National Health and Nutrition Examination Survey (KNHANES) from 2010 to 2014 (n = 2,198,964), to investigate current Korean food and nutrients intake. These datasets targeted sub-populations (about ten thousand people) categorized into children (1–11 yrs), youths (12–18 yrs), and adults (≥19 yrs). Moreover, cohort data consisted of a total of seven (n = 321,100) studies of groups in local communities, rural areas, and cities, as well as of twins and their families.

Second, there were a total of five food composition datasets (n = 88,248), including the Korean National standard Food Composition Table (KNFCT) and national monitoring datasets for processed foods.

Third, there were three HBGV datasets (n = 106), including the Korean Daily Recommended Intake (KDRI).

All core datasets collected are shown in Table 1.

2.2. Compilation of core datasets

2.2.1. Compilation processes

All core datasets were collected in the form of raw-data with a coding book that explains the variable name, type, description and coding specification. The datasets had different formats (e.g., Microsoft Excel, Access or SAS). The compilation processes for building an integrated database were as follows: 1) import the raw-data into the Relational Database Management System (RDBMS), 2) code standardization, 3) encode and transfer raw-data into an integrated database. RDBMS used Oracle 11 g (Oracle Inc.).

2.2.2. (Step 1) Import raw-data into RDBMS

This step was designed to convert multiple core dataset files with different formats into only one Oracle database file. To accomplish this, all raw-data were imported into the Oracle server without transforming their original contents and coding specifications.

2.2.3. (Step 2) Code standardization

This step was conducted to remove redundancy and interconnect different datasets by standardizing the different code values of variables with the same meaning (i.e., food and nutrient codes) uniformly. The collisions by code incompatibility shown in Fig. 2 occurred not only in different datasets, but also within the same datasets (e.g., KNHANES) because of different production terms among datasets.

Initially, code standardization for food was executed using the 5th (2010–2012) and 6th (2013–2014) KNHANES datasets because they contained the abundant most food codes. Some codes were duplicated by errors and discovered in a certain year based on emerging food due to changing eating habits. To solve these problems, we applied the technical database normalization process and executed the total inspection for all food descriptions manually. Finally, the representative, time series styled and standardized food codes, including hierarchical structure with the 1st, 2nd, and 3rd level, were elicited.

To standardize the nutrient codes, we used three food composition datasets: FANTASY, KNFCT, and NSC. Finally, the representative and standardized nutrient codes, including the hierarchical structure with the 1st category (macro/micronutrients) and 2nd category (macro/