USDA's Expanded Flavonoid Database for the Assessment of Dietary Intakes

Documentation and User Guide

Prepared by

Seema Bhagwat, David B. Haytowitz and Shirley Wasswa-Kintu

Nutrient Data Laboratory
Beltsville Human Nutrition Research Center
Agricultural Research Service
U.S. Department of Agriculture

September 2014

U.S. Department of Agriculture
Agricultural Research Service
Beltsville Human Nutrition Research Center
Nutrient Data Laboratory
10300 Baltimore Avenue
Building 005, Room 107, BARC-West
Beltsville, Maryland 20705
Tel. 301-504-0630, FAX: 301-504-0632
E-Mail: ndlinfo@ars.usda.gov
Web site: http://www.ars.usda.gov/nutrientdata

Table of Contents

Suggested Citation: i
Acknowledgements i
Documentation 1
Developing the Expanded Flavonoids Database 1
  Matching analytical values 2
  Calculating flavonoid values 3
  Calculating flavonoid values for multi-ingredient foods 4
Limitations 6
Format of the Database 6
  Food Description File 7
  Food Group Description File 7
  Flavonoid Data File 7
  Nutrient Definition File 8
  Source Code File 9
  Data Derivation Code Description File 9
References: 9

Suggested Citation:

Bhagwat S., Haytowitz, D.B., and Wasswa-Kintu S. 2014. USDA's Expanded Flavonoid Database for the assessment of Dietary Intakes. U.S. Department of Agriculture, Agricultural Research Service. Nutrient Data Laboratory Home Page: http://www.ars.usda.gov/nutrientdata

Acknowledgements

The authors wish to acknowledge Pamela Pehrsson (NDL, Research Leader) and Joanne M. Holden (NDL Research Leader, Retired) for their contributions and support during the development of this database. The authors also gratefully acknowledge the constructive contributions of the peer reviewers: Jaspreet K.C. Ahuja, Nutrient Data Laboratory, USDA-ARS, and Gary Beecher, Food Composition and Methods Development Laboratory, USDA-ARS (Retired)

USDA's Expanded Flavonoid Database for the Assessment of Dietary Intakes

Documentation

The potential beneficial health effects of dietary flavonoids continue to interest the scientific community in associating the flavonoid intakes and certain chronic diseases (Schroeter et al., 2010; Erdman et al., 2007). Many knowledge gaps exist in the study of flavonoids, including the lack of a comprehensive food composition database, making assessment of population intake of flavonoids difficult. Scientists at the Nutrient Data Laboratory (NDL) addressed these gaps in the database by constructing a new flavonoids database based on the analytical values from the "USDA Database for the Flavonoid Content of Selected Foods, Release 3.1", 2013" (FDB 3.1) and the "USDA Database for the Isoflavone Content of Selected Foods, Release 2.0, 2008", (IDB 2.0). This database provides values for the subset of approximately 2,900 foods from the National Nutrient Database for Standard Reference, Release 22 (SR22) used in the development of the Food and Nutrient Database for Dietary Studies (FNDDS), 4.1. Using the analytical values for 26 flavonoid compounds for about 500 food items in FDB 3.1 and three isoflavones for over 550 food items in IDB 2.0 as the foundation, flavonoid profiles for the ~2900 food items were generated for the total of 29 flavonoid compounds to develop the Expanded Flavonoid Database for the Assessment of Dietary Intakes (FDB-EXP). Unlike FDB 3.1 and IDB 2.0, FDB-EXP reports values for all 29 flavonoid compounds for every food and are a combination of analytical and calculated values. All the analytical values in the FDB 3.1 and IDB 2.0 were retained in the FDB-EXP. The unit of measure for the flavonoid compounds in the expanded database is mg/100g edible portion on fresh weight basis. The NDL has provided this data set to the Food Surveys Research Group (FSRG) to develop flavonoid values for over 7,000 foods in FNDDS 4.1, which will be used to estimate flavonoid intakes of the U.S. population. This work was supported by the Office of Dietary Supplements (ODS), NIH, DHHS.

Dietary flavonoids include flavonoids in six subclasses (Table 1) and polymeric flavanols (proanthocyanidins, theaflavins and thearubigins). They are stress induced, secondary plant metabolites and are present in fruits, vegetables, nuts and some grains. Isoflavones occur mainly in soybeans and soybean products. A complete flavonoid profile, made up of the 29 selected commonly occurring dietary flavonoids, was developed for each food item in the SR subset. The polymeric flavanols, proanthocyanidins, were not profiled for this subset at this time.

Developing the Expanded Flavonoids Database

Populating over 2,900 food items of the SR subset with full flavonoid profiles consisting of 29 commonly occurring dietary flavonoid compounds from six flavonoid subclasses required identifying food categories based on flavonoid contents, determining the approach for calculating flavonoid values, and calculating values where necessary by considering various factors that affect the values. NDL scientists identified those food groups which are thought not to contain flavonoids as well as food groups which may

contain flavonoid compounds based on FDB 3.1 and IDB 2.0. Multi-ingredient foods that may have flavonoid containing ingredients were also identified.

Table 1. Six Subclasses of flavonoids and selected compounds

Flavonoid Subclass	Flavonoid Compounds
Flavonols	Isorhamnetin, Kaempferol, Myricetin, Quercetin
Flavones	Apigenin, Luteolin
Flavanones	Eriodictyol, Hesperetin, Naringenin
Flavan-3-ols	Catechin, Epicatechin, Epicatechin 3-gallate, Epigallocatechin, Epigallocatechin 3-gallate, Gallocatechin, Theaflavin, Theaflavin 3-gallate, Theaflavin 3'-gallate, Theaflavin 3,3'-digallate, Thearubigins
Anthocyanidins:	Cyanidin, Delphinidin, Malvidin, Pelargonidin, Peonidin, Petunidin
Isoflavones	Daidzein, Genistein, Glycitein

Assigning "zero" values

Animal based food groups (finfish, shellfish, poultry, beef, dairy, fats, oils, sausages, luncheon meats, and pork items) generally were not expected to contain flavonoids and a value of zero was assigned for all flavonoids in these food items. Based on limited data in FDB 3.1 and IDB 2.0, and an extensive review of literature, flavonoid data was not found for foods in these groups, except a very small amount in eggs. In the absence of analytical data, multi-component foods derived from these food groups were also assigned zero values for all 29 flavonoid compounds if other ingredients were not expected to contain flavonoids. Additional research is needed to obtain information in order to fully characterize the flavonoids content of these foods.

Only one or two subclasses of flavonoids are predominant in most food items, e.g. flavanones are a major subclass in citrus fruits. Therefore zero values were assigned to flavonoid compounds in the subclasses that were not expected to be present in a particular food or food group.

Matching analytical values

Food groups expected to contain one or more flavonoid subclasses included fruits, fruit juices, vegetables, vegetable products, spices, herbs, nuts, seeds, and legumes, especially soybeans (Middleton 1996). To determine complete flavonoid profiles for all plant types, specific food descriptions and analytical values from FDB 3.1 and IDB 2 were matched with the corresponding Nutrient Data Bank numbers (NDB) in SR for the foods in the SR Subset and all the available analytical values for the matched food items in FDB 3.1 and IDB 2.0 were retained in the FDB-EXP.

Values for instant tea powders provided to the FSRG were based on the data provided by Unilever Lipton Company in 2002. The weight for added flavoring and /or artificial sweetener like saccharin was disregarded, i.e. same values were used for unsweetened/unflavored and sweetened with saccharin/flavored instant and prepared teas.

Calculating flavonoid values

After assigning available analytical values to those foods which were direct matches by food descriptions in FDB 3.1 and IDB 2.0, various techniques were then employed to calculate values for the remaining compounds and/or foods according to the procedures described by Schakel et al., (1997).

The various procedures used to calculate values where analytical values were unavailable are described below.

Moisture Adjustment: For plant-based foods that underwent a moisture adjustment due to cooking, drying or dilution, values for full flavonoid profiles were calculated from another form of the same food or from a similar food. A factor based on the change in total solids was applied (e.g. raw to cooked asparagus or fresh to dried basil).

Retention Factors: To calculate values for compounds after cooking/processing of the foods retention factors were determined by consulting scientists at the Food Composition and Methods Development Laboratory (FCMDL) of ARS/USDA as well as from limited literature sources. As a result, retention factors of 85% for flavonols, flavanols, flavanones and flavones and 50% for anthocyanidins (because of their heat labile characteristics) were established (FCMDL, personal communication, 2012). The application of dry heat (i.e., baking) was considered to have negligible effects on flavonoid losses (James Harnly, FCMDL, personal communication, 2012). Therefore no retention factors were applied when drying process was used, e.g. for fresh to dried fruits or fresh to dried herbs. If the analytical values were available for cooked/processed foods, these values were kept in preference to those calculated by applying retention factors.

In the European Prospective Investigation into Cancer and Nutrition (EPIC) study Zamora-Ross et al., (2011a and 2011b) used FDB release 2.1 (2007) and Phenol-Explorer, Release 2 (Neveu et al; 2010) to estimate intakes of flavanols, flavones, flavanones and anthocyanidins using retention factors of 70%, 35%, and 25% for fried, microwaved, and boiled foods. These factors are different from the NDL factors and may produce discrepancies in the values for cooked foods in these two databases.

NDL in collaboration with the FCMDL is conducting a study to generate retention factors for individual flavonoid compounds reported in these databases after different cooking methods. Recently Phenol-Explorer database has published retention factors for polyphenols including flavonoids using data from the published literature (Rothwell et al., 2013). These factors were not available at the time FDB-EXP was developed.

Cooked or canned plant based foods without full flavonoid profiles were calculated from values for raw forms of the same food using estimated retention factors to account for the loss of flavonoids during processing.

Retention factors were not needed for isoflavones, as analytical values were available for most of the raw and cooked/processed foods that contain isoflavones.

Food Yield Factors: Food yield factors were also applied to account for food processing effects in instances when values were available only for a different form of the same food. For example, for canned foods, a yield factor of approximately 0.67 was applied to the raw form to adjust for yields of solid foods after draining liquids (USDA, 1975).

Substitution with similar foods: Values from a similar food were substituted for the food or compound considering botanical origins or other similarities.

Market Share: Other factors such as market share data were considered when estimating flavonoid values for some items. Market share proportions of 47% red wine, 40% white wine, and 13% blush wine were reported by ACNielsen supermarket data, U.S. Department of Commerce, 2011. For the lack of data for blush wines (only one source), a generic profile for table wine was created by using 50% of red and 50% of white wine values. In the development of FDB 3.1, some food items (e.g., red and green grapes) were assigned more specific provisional NDB numbers to differentiate them from each other on the basis of flavonoid content, although these foods have a single NDB number in SR. A generic value for all grapes, corresponding to the NDB number used in SR, was developed by using both the red (50%) and green (50%) grape values.

Generic Profiles: A generic profile was prepared for common leafy vegetables using values from FDB 3.1 to estimate values for other leafy vegetables where values were not available. A generic profile for fruit was also prepared to be used for less common fruits when similar fruits were not available.

Other sources: Values from other databases such as Phenol-Explorer, Release 2 (Neveu et al; 2010) were used for some foods not available in the FDB 3.1. For instance kaempferol values for most nuts were obtained from Phenol-Explorer.

Calculating flavonoid values for multi-ingredient foods

Multi-ingredient foods with one or more ingredients of plant origin such as baby foods, soups, breakfast cereals, beverages, may also contain some flavonoids, depending on amounts of the ingredients and effects of processing. However, few multi-ingredient food items were included in FDB 3.1 because of the lack of analytical data. For multi-ingredient foods, formulations developed by NDL scientists (Haytowitz et al., 2009) were used to estimate percentages of flavonoid-containing ingredients. An example of this process is also included in Table 2.

Table 2: Examples of techniques used to calculate values where analytical data was not available

Calculation Technique	NDB No.	Food Description	Source NDB No. for Calculation	Source Food Description	Flavonoid Subclass	Factor Used
Retention and moisture Adjustment	11162	Collard, cooked, boiled, drained	11161	Collard, raw	Flavonol (Isorhamnetin)	0.85 flavonol retention and 0.86 moisture ratio
Yield Adjustment	09374	Pear, canned, heavy syrup, drained	09252	Pears, raw	All classes	0.67
Substitution with similar food	09110	Mulberries, raw	09042	Blackberries, raw	Flavan 3-ols, Flavanones and Anthocyanidins	N/A
Moisture Adjustment	02031	Spice, pepper, red or cayenne, dried.	99369	Cayenne pepper, raw	All classes	11.80
% Plant-based ingredient formulation for a mixed dish	06025	Soup, chicken vegetable, canned, condensed	12% (11357)	Potatoes, white, flesh and skin, baked	All classes	0.12
			23% (11125)	Carrots, cooked, boiled, drained, w/o salt	All classes	0.23
			12% (11530)	Tomatoes, red, ripe, cooked	All classes	0.12

In general, flavonoid values for each basic ingredient were used to calculate flavonoid contents of the multi-ingredient foods only if the individual ingredient contributed more ≥ 5% to that particular food. Cocoa powder, regular or alkalinized, soy protein isolate and soy flour were the exceptions due to their observed high contents of flavan-3-ols and isoflavones respectively per unit weight. Multi-ingredient foods containing less than five percent of each plant-based ingredient were estimated to have no flavonoids present in that food. Some foods (bakery products or tomato chili sauce) may have unexpectedly high flavonoid contents due to the use of ingredients such as soy protein isolates. Whenever orange juice was one of the ingredients (≥5%), e.g. in baby food juices or citrus juice drinks, we used values of “orange juice, chilled, includes from concentrate (NDB No. 09209)” to calculate flavanone values for these foods for consistency.

Limitations

The database is fully documented regarding how values were calculated (derivation codes used in the SR) and if appropriate, the NDB number of the source food from which the values were obtained.

Values for the 29 different flavonoids were obtained from a variety of sources, as were the samples analyzed. The data reflect samples obtained over time from many sources representing different cultivars, growing locations, harvesting and processing techniques, laboratories, and analytical methods. Therefore, in a comparison of different forms of a particular food (e.g. raw and cooked), differences in flavonoid content may not measure precisely the effect of processing or preparation methods. For example, the value for quercetin in raw onions (NDB no. 11282) was 20.30 mg/100g, but in the cooked onions (NDB no. 11283) was 24.36 mg/100g. If a value in the cooked food is more or less than in the raw food, it does not necessarily mean that the particular flavonoid was increased or reduced because of cooking. This kind of comparison is valid only when paired raw and cooked samples are analyzed to estimate values for these forms.

Total flavonoid values measured by the colorimetric methods such as Folin-Ciocalteau are not included because of the lack of specificity of the methods.

Summary

Plant foods, such as, fruits, vegetables, legumes, and a few grains are the main sources of flavonoids. Animal foods are generally not expected to contain flavonoids. Only one or two subclasses of flavonoids are predominant in most foods, e.g. flavanones in citrus. Generally, researchers focus on analyzing the predominant compounds only. Therefore values for the full profile of all 29 flavonoid compounds are not available for many foods. It is necessary to create full profiles to estimate flavonoid intakes for NHANES.

This expanded database contains data for 29 flavonoids in six classes for 2,926 food items from SR 22 which provides the basis for the FNDDS 4.1. Food descriptions and analytical data matched 367 food items which were taken from 913 unique food items contained in FDB 3.1 and IDB 2.0 databases. Zero values were assumed for ~73% of the total 87,780 values in the database. Of the non-zero values, ~3% were analytical and ~24% were calculated using the procedures described above.

This data set was provided to FSRG, USDA to develop flavonoid values for more than 7,000 foods in FNDDS 4.1, which will be used to estimate flavonoid intakes of the U.S. population, based on the What We Eat in America component of NHANES 2007-08. This will allow researchers to use comprehensive data on flavonoid intakes to assess the impact of these bioactive components on human health.

Format of the Database

For the convenience of the user, the Expanded Flavonoid Database for the Assessment of Dietary Intakes has been imported into a Microsoft® Access 2007 database

(FDB-EXP.accdb). This database follows the same structure as that used for SR thus allowing users to access the database in a form compatible with other programs. It can be linked to SR to obtain additional information on the food items. Links indicating the relationships among the files are presented with each file.

The tables and fields in the Microsoft® Access 2007 database are as follows:

Food Description File (file name = FOOD_DES). This file (Table 3) contains the descriptions of the food items. For those items in the SR additional information (e.g., common names, percentage, and description of refuse) can be obtained by linking this table to the corresponding table in SR.

- Links to the Food Group Description file by FdGrp_Cd
- Links to the Flavonoid Data file by NDB No.
- Links to the Flavonoid Detail file by NDB No.

Table 3.—Food Description File Format

Field Name	Type	Description
NDB_No*	A 5	5-Digit Nutrient Databank number that uniquely identifies a food item. Foods in the USDA
FdGrp_Cd	A 4	4-digit code indicating food group to which the food item belongs
Long_Desc	A 200	Description of the food item

*Primary key for the food description file

Food Group Description File (file name = FD_GROUP). This file (Table 4) contains a list of food groups used in the flavonoid database and their descriptions.

- Links to the Food Description file by FdGrp_Cd

Table 4.—Food Group Description File Format

Field Name	Type	Description
FdGrp_Cd*	A 4	4-digit code identifying a food group. Only the first two digits are currently assigned. All of the food groups in SR are not used in the flavonoid database.
FdGrp_Desc	A 60	Name of food group

* Primary key for the Food Group Description file.

Flavonoid Data File (file name = FLAV_DAT)**.** This file (Table 5) contains the flavonoid values. Information about analytical values, such as statistical information, confidence codes, and sources of data can be obtained from the corresponding database—FDB 3.1 and IDB 2.0..

- Links to the Food Description file by NDB No.
- Links to the Nutrient Definition file by Nutr. No.
- Links to the Source Code file by Src_Cd
- Links to the Data Derivation Code Description file by Deriv_Cd

Table 5.—Flavonoid Data File Format

Field Name	Type	Description
NDB No.*	A 5	5-Digit Nutrient Databank number
Nutr_No*	A 3	Unique 3-digit identifier code for each flavonoid compound
Flav_Val	N 10.2	The flavonoid mean value (mg/100 g) edible portion (analytical or calculated)
Src_Cd	A 2	Code indicating type of data.
Deriv_Cd	A 4	Data Derivation Code giving specific information on how the value is determined
Ref_NDB_No	A 5	NDB number of the item used to calculate a missing value.

* Primary keys for Flavonoid Data file.

Nutrient Definition File (file name = NUTR_DEF). This file (Table 6) contains the nutrient number and the description of the flavonoids.

- Links to the Nutrient Data file by Nutr_No.

Table 6.—Nutrient Definition File Format

Field Name	Type	Description
Nutr_No*	A 3	Unique 3-digit identifier code for each flavonoid
Flav_Class		The subclass of flavonoids to which the individual flavonoid belongs
Description	A 60	Name of the flavonoid
Tagname	A 20	International Network of Food Data Systems (INFOODS) Tagnames.† A unique abbreviation for a nutrient/food component developed by INFOODS to aid in the interchange of data.
Unit	A 7	Units of measure (e.g. mg)
Decimal	A 1	Number of decimal places to which the flavonoid values is displayed.

* Primary key for Nutrient Definition file.
† INFOODS, 2014.

Source Code File (file name = SRC_CD). This file (Table 7) contains codes indicating the type of data (analytical, calculated, assumed zero, and so on) in the Nutrient Data file. To improve the usability of the database and to provide values for the FNDDS, NDL staff imputed nutrient values for a number of proximate components, total dietary fiber, total sugar, and vitamin and mineral values.

- Links to the Nutrient Data file by Src_Cd

Table 7.—Source Code File Format

Field name	Type	Blank	Description
Src_Cd*	A 2	N	2-digit code.
SrcCd_Desc	A 60	N	Description of source code that identifies the type of nutrient data.

* Primary key for the Source Code file.

Data Derivation Code Description File (file name = DERIV_CD). This file (Table 8) provides information on how the nutrient values were determined. The file contains the derivation codes and their descriptions.

- Links to the Nutrient Data file by Deriv_Cd

Table 8.—Data Derivation Code File Format

Field name	Type	Blank	Description
Deriv_Cd*	A 4	N	Derivation Code.
Deriv_Desc	A 120	N	Description of derivation code giving specific information on how the value was determined.

* Primary key for the Data Derivation Code file.

References:

Bhagwat S, Haytowitz DB, Holden JM. USDA Database for the Flavonoid Content of Selected Foods, Release 3.1. [online] Nutrient Data Laboratory. Beltsville (MD): Beltsville Human Nutrition Research Center, Agriculture Research Service, USDA; 2013 [cited 2014 September 4]. Available from: http://www.ars.usda.gov/Services/docs.htm?docid=6231

Erdman Jr, JW Balentine D, Arab L, Beecher G, Dwyer JT, Folts J, Harnly J, Hollman P, Keen CL, Mazza G, Messina M, Scalbert A, Vita J, Williamson G, and Burrowes J.

Flavonoids and heart health: Proceedings of the ILSI North America flavonoids workshop, May 31-June 1, 2005, Washington, DC. J. Nutr. 2007;137:718S-737S.

Haytowitz DB, Lemar LE, and Pehrsson PR. USDA's Nutrient Databank System – A tool for handling data from diverse sources. J. Food Comp. Anal. 2009;22:433-441.

Food and Agricultural Association of the United Nations. INFOODS. Tagnames for Food Components. 2014; [cited 2014 September 4]. Available from: http://www.fao.org/infoods/infoods/standards-guidelines/food-component-identifiers-tagnames/en/.

Middleton E Jr. Biological properties of plant Flavonoids: an overview. Int J Pharm. 1996; 34:344-348.

Neveu V, Perez-Jiménez J, Vos F, Crespy V, du Chaffaut L, Mennen L, Knox C, Eisner R, Cruz J, Wishart D, Scalbert A. Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. *Database; 2010* [cited 2014 September 4]. Available from: *http://database.oxfordjournals.org/content/2010/bap024.short docid=bap024*

Rothwell JA, Perez-Jimenez J, Neveu V, Medina-Ramón A, M'Hiri N, Garcia-Lobato P, Manach C, Knox C, Eisner R, Wishart DS, and Scalbert A. Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. Database. 2013; [cited 2014 September 4] docid=bat070

Schakel SF, Buzzard IM, Gebhardt SE. Procedures for estimating nutrient values for food composition databases. J. Food Comp. Anal. 1997;10, 102-114.

Schroeter H, Heiss C, Spencer JPE, Keen C, Lupton JR, and Schmitz HH. Recommending flavanols and procyanidins for cardiovascular health: Current knowledge and future needs. Mol. Aspects. Med. 2010; 31, 546-557.

U.S. Department of Agriculture, Agricultural Research Service. USDA Database for the Isoflavone Content of Selected Foods, Release 2.0. 2008. [cited 2014 September 4]. Available from: http://www.ars.usda.gov/Services/docs.htm?docid=6382

U.S. Department of Commerce. U.S. Wine Industry – 2011. 2011. [cited 2014 September 4]. Available from: http://www.ita.doc.gov/td/ocg/wine2011.pdf

U.S. Department of Agriculture, Agricultural Research Service. Food Yields, summarized by different stages of preparation. Agriculture Handbook No. 102;1975.[cited 2014 September 4] Available from: http://www.ars.usda.gov/SP2UserFiles/Place/12354500/Data/Classics/ah102.pdf