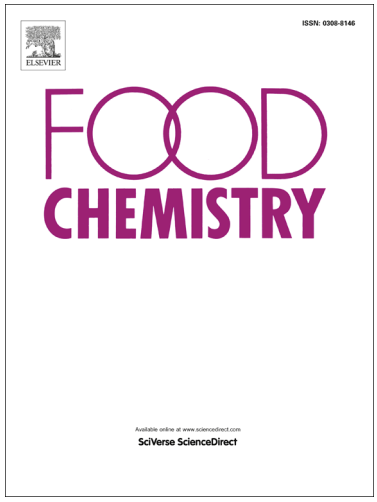


A systematic review of food composition tools used for determining dietary polyphenol intake in estimated intake studies

Yasmine Probst, Vivienne Guan, Katherine Kent

PII: S0308-8146(16)31834-9
DOI: http://dx.doi.org/10.1016/j.foodchem.2016.11.010
Reference: FOCH 20143

To appear in: *Food Chemistry*

Received Date: 11 May 2016
Revised Date: 18 October 2016
Accepted Date: 2 November 2016

Please cite this article as: Probst, Y., Guan, V., Kent, K., A systematic review of food composition tools used for determining dietary polyphenol intake in estimated intake studies, *Food Chemistry* (2016), doi: http://dx.doi.org/10.1016/j.foodchem.2016.11.010

A systematic review of food composition tools used for determining dietary polyphenol intake in estimated intake studies

Running title:

Tools for determining dietary polyphenol intake

Authors:

Yasmine Probst[a,b*], Vivienne Guan[c] and Katherine Kent[b,d]

Affiliations:

[a]School of Medicine, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong NSW 2522 Australia, Email: yasmine@uow.edu.au

[b]Illawarra Health and Medical Research Institute, Wollongong NSW 2522

[c]School of Medicine, Faculty of Science, Medicine and Health, Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong NSW 2522 Australia, Email: xg885@uowmail.edu.au

[d]School of Medicine, Faculty of Science, Medicine and HealthUniversity of Wollongong, Wollongong NSW 2522 Australia, Email: kc582@uowmail.edu.au

* Corresponding author for all correspondence and stages of review and publication.

Abstract:

Translating food intake data into phytochemical outcomes is a crucial step in investigating potential health benefits. The aim of this review was to examine the tools for determining dietary-derived polyphenol intakes for estimated intake studies. Published studies from 2004 to 2014 reporting polyphenol food composition information were sourced with 157 studies included. Six polyphenol subclasses were identified. One quarter of studies (n=39) reported total flavonoids intake with 27% reporting individual flavonoid compounds. Assessing multiple compounds was common with approximately 10% of studies assessing seven (n=13), six (n=12) and five (n=14) subclasses of polyphenol. There was no pattern between reported flavonoids compounds and subclass studied. Approximately 60% of studies relied on publicly accessible food composition data to estimate dietary polyphenols intake with 33% using two or more tools. This review highlights the importance of publicly accessible composition databases for estimating polyphenol intake and provides a reference for tools available globally.

Key words: Phytochemical; polyphenol; food composition data; systematic literature review; dietary assessment; observational studies

Chemical compounds studied in this article:

Anthocyanin (CID 145858), 3-Flavanol (CID 12318031), Flavanone (CID 10251), Flavone (CID 10680), Flavonol (CID 11349), Isoflavone (CID 72304), Daidzein (CID 5281708), Genistein (CID 5280961), Lignan (CID 159949)

1.0 Introduction:

The evidence underpinning the Australian Dietary Guidelines specifically relates the consumption of core plant based food groups (Beecher, 2003) - fruit, vegetables and grains - to phytochemicals (carotenoids, flavonoids and isoflavonoids, polyphenols, xanthin etc.) consumption (Department of Health and Ageing & National Health and Medical Research Council, 2011). Despite this, the 2011-13 Australian Health Survey indicates that only 5.6% of Australian adults achieved the recommended two and five servings of fruit and vegetables, respectively (Australian Bureau of Statistics, 2012), a pattern that has persisted for many decades (Australian Institute of Health and Welfare, 2013; Magarey, McKean, & Daniels, 2006). In parallel, research also is looking to define how plant based foods truly impact on health outcomes (Tapsell, Dunning, Warensjo, Lyons-Wall, & Dehlsen, 2014). Consumption of total phytochemical intake is consistently linked with protection against chronic diseases (Knekt et al., 2002), including cardiovascular disease (Hooper et al., 2008), cancer (Park & Pezzuto, 2012) and neurodegenerative diseases (Commenges et al., 2000).

Application of food composition data remains at the forefront of dietetic practice (Dietitians Association of Australia, 2010), though translation from the nutrient to the food information needs to be strengthened to better support public health messages. In order to associate phytochemical consumption with positive health outcomes, a fundamental step is to accurately estimate dietary phytochemical intake.

Despite the first estimations of phytochemical intake at a population level being reported more than a decade ago, the numerous methods employed have evident flaws (Dwyer & Peterson, 2002). Dietary phytochemical intake is difficult to quantify and

consequently numerous methods have been developed for application in various settings. With the absence of a gold standard approach, the methods utilised include various techniques within the fields of dietary assessment and biomarker analyses. Specifically for translation to occur at a nutrient level or at the grouped food level to create advice strategies, an up-to-date and geographically appropriate food composition database is required. Translating food data specifically to phytochemical intakes is further complicated due the number of phytochemicals found intrinsically in foods, their bioavailability when consumed and their interactions with other foods or nutrients when consumed as part of the whole diet.

The limitations associated with current methods hinder the interpretation of research outcomes that associate dietary phytochemical intake and specific health outcomes. An evaluation and comparison of the tools to measure phytochemical intake is imperative to interpret current findings across the literature and to provide recommendations for methods to apply in future research.

As phytochemicals is the term used to group a vast range of chemical compounds which are hieratically grouped into classes and subclasses, Unlike other known nutrients in foods, the complexity and variability must also be carefully considered. Traditional methods of dietary assessment require a recall or documentation of food intake from a given time period in either a prospective or retrospective manner. To determine the nutrient composition of either individual or group intakes, this dietary intake data must have tools applied to it to allow a food to nutrient translation to occur. These tools may food composition databases, limited for phytochemicals, or relate directly to the intake data or the use of known biomarkers detected in the

plasma, urine or faecal samples of the person giving the recall to confirm the plausibility of the intake data that has been provided.

Dietary assessment of phytochemical intake

The most common method of estimating phytochemical intake at a population level relies on dietary assessment of intake. Generally, assessment of usual diet may be performed using repeated 24-hour diet recalls, diet history interview or food frequency questionnaires. These methods are then cross-referenced with a phytochemical food composition database. However, there are very few phytochemical specific food composition databases that exist globally. Aside from the limitations inherent to each dietary assessment method, there are several well documented problems associated with utilising food composition databases not specific to the geographic area to assign phytochemical content to selected foods, resulting in large variations in estimates of intake (Chun, Lee, Wang, Vance, & Song, 2012).

Firstly, estimation of dietary phytochemical intake is only as comprehensive as the composition database utilised. If, for example, a composition database does not have an extensive list of foods and the phytochemical content of a food in an individual's diet cannot be assigned or matched to its closes equivalent, and in turn an individual's intake will be underestimated. This is particularly challenging when analysing food intake data from a country that does not have a specific composition database for that population. Secondly, the phytochemical content of specific foods is highly variable and largely influenced by a foods growth, harvesting and processing conditions. A phytochemical food composition database is unable to account for this variability and

can only provide an estimate for each food consumed. Lastly, estimating dietary phytochemical intake through dietary assessment is unable to account for the high intra-individual variation associated with phytochemical metabolism and absorption, which is influenced by factors other than intake, such as bioavailability and genetic factors. Until the bioavailability of all phytochemicals are understood and the individual variations in metabolism are accounted for, estimations of phytochemical intake and their correlation with health outcomes should be interpreted with caution.

Biomarker analyses for phytochemical intake

Dietary phytochemical intake can be determined by quantifying biomarkers which include intact phytochemicals and their derivatives (eg. phenolic acids) found in plasma, urine and faecal water. Many methods of measuring phytochemical biomarkers in human biological samples exist, with no standardised protocol of how to perform this analysis. Consequently researchers must develop and validate their own methods, limiting the ability to compare studies that have used different methods to measure certain biomarkers. Generally, laboratories use chromatography and spectrometry to quantify the biomarker of interest. However, there many thousands of phytochemicals identified and after consumption they are quickly and extensively metabolised into various metabolites. Consequently, there are thousands of potential biomarkers and there is no consensus around which phytochemicals or metabolites are indicative of total dietary intake.

More recently the use of metabolomics, the analysis of all metabolites contained in a given biofluid at a given time, in combination with pattern recognition analyses and advancements in analytical have been employed to search for relevant biomarkers.

This approach provides improved specificity though is also limited by the biological measures it can address (Monteiro, Carvalho, Bastos, & Guedes de Pinho, 2013; O'Gorman, Gibbons, & Brennan, 2013). A recent metabolomics-based study into biomarkers of high and low flavonoid intakes from fruit and vegetables identified abscisic acid glucuronide for the first time in relation to low flavonoid dietary intakes while confirming phenolic acids and their derivatives in relation to high intakes (Ulaszewska et al., 2016), demonstrating that biomarkers may need to be suited to both the component being metabolized and the context in which is it being considered.

In addition, it is currently unknown which biological sample (plasma, urine or faecal water) should be selected and research suggests each may be indicative of different consumption patterns. Previous research shows urinary biomarkers may be more reflective of short-term intake (Radtke, Linseisen, & Wolfram, 2002). The phytochemical content in fasting plasma or faecal water samples seems to be a suitable biomarker of short-term intake and a possible biomarker of the medium-term intake (Radtke et al., 2002). However, biomarkers of long-term intake are not yet identified and may be unlikely due to the short half-lives of dietary phytochemicals in vivo. Most of the biomarker analyses are expensive and often cannot be performed as part of large epidemiological studies (Yokota, Miyazaki, & Ito, 2010). Future research needs to focus on identifying specific biomarkers of phytochemical intake and confirm the best methods in which to quantify these biomarkers in biological specimens, to inform population research.

With no gold standard method for measuring phytochemical intake, it is unclear which method for measuring or estimating dietary phytochemical intake is most useful. To improve methodological quality of research, a clear understanding of appropriate methods for measuring phytochemical intake is required. This review aims to provide an overview of available strategies for estimating dietary phytochemical intake and to provide an important resource for researchers.

2.0 Materials and methods:

This review is registered with the International Prospective Register of Systematic Reviews (PROSPERO) under the registration number #CRD42014015607. The structure of this review followed the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009) with the following question used to guide the literature search: "What tools are used to determine intake of dietary phytochemicals?" Due to the wide range of phytochemicals compounds, this review will focus on the most studied sub-class polyphenols (Figure 1).

Published studies from January 2004 through to November 2014 reporting food composition information for polyphenols were sourced. The tools for dietary polyphenol intake data were examined, as were patterns of results among polyphenol subclasses and the use of different tools. Due to the wider variability related to biomarkers of intake and its emerging evidence base, this review will only focus on the translation of food to nutrient intakes via food composition related tools (databases, tables or other published works).

2.1 Search strategy

The search aimed to find both published and unpublished studies through electronic databases, the Internet and hand searching of reference lists. Key terms were used in the following truncated form: Phenolic acid OR Flavonoid* OR Flavanol* OR Stilbene* Or Lignans OR Isoflavone* OR Anthocyanin* Or Flavanone* OR Flavonol* OR Flavone* OR Catechin OR Ellagic acid OR Genistain OR Polyphenol* OR flavan-3-ol*. The first stage included searches conducted using the Web of Science and Scopus scientific databases, while a second stage included searching in the following Internet sites using the key terms to capture Australian databases.

- Food Standards Australian New Zealand http://www.foodstandards.gov.au
- National Health & Medical Research Council www.nhmrc.gov.au
- Australian Institute of Health & Welfare www.aihw.gov.au
- World Health Organization http://www.who.int/en/
- Australian Bureau of Statistics http://www.abs.gov.au

2.2 Eligibility criteria

2.2.1 Types of studies

This review included analytical epidemiological study designs including prospective and retrospective cohort, case-control and cross-sectional studies. Our preliminary data extraction shown randomised, non-randomised food-based trials and crossover food-based trials tended to not estimate polyphenols intake from the whole diet by using tools. Only studies published in English language were considered for inclusion due to a lack of translation resources.

2.2.2 Types of data

The main data extracted for this review were the reported polyphenol types and tools used for estimating dietary polyphenols intake. The details for the use of tools for dietary polyphenol intake data were examined, as were patterns of result combinations among polyphenols between different tools used.

2.2.3 Types of methods

Studies reporting data for whole food based polyphenols outcomes in relation to a health condition were included. Studies reporting the use of a tool for the translation of food intake to polyphenols data, such as a food composition database were also included.

Studies that did not measure whole foods or whole of diet based polyphenols were excluded, this included studies related to the use of supplements, encapsulated polyphenol extracts, extract from herbal sources and purified or modified version of polyphenols. Studies related to bioavailability or mechanistic feeding trials were also excluded. The polyphenol-containing foods considered needed to be commercial available or publicly accessible by the general population.

2.2.4 Types of outcome measures – Primary outcomes

The primary outcome of the systematic review was the summary of reported polyphenol types and tools used for estimating dietary polyphenols intake.

2.3 Data collection and analysis

2.3.1 Selection of studies

The review was structured and reported according to the PRISMA. One review author (YP) conducted the literature search in the specified scientific databases. Two additional review authors (VG and KK) independently assessed and compared potential studies identified by the search strategy for inclusion. Resolution of any disagreements occurred through discussion and required a consensus outcome. Where consensus could not be reached a third researcher (YP) was consulted.

Articles identified by database searches were assessed for relevance to the review based on the title and abstract (Table 1). For those meeting the inclusion criteria, the full text publication were retrieved and assessed for relevance to the review criteria.

Table 1: Overall inclusion and exclusion criteria for screening

Section	**Criteria**	**Include if**
Language	Study reported in English	Yes
Design	Prospective or retrospective cohort or,	Yes
	Case-control, cross-sectional study.	
	Case reports, reviews, editorials, letter to the editor, qualitative research and short communication	No

Population	Adults aged >18 years	Yes
	Animal study or study including persons <18 years	No
Content	Study examines the tools for estimating dietary-derived polyphenols intake.	Yes
	Study examines encapsulated phytochemicals, extract from herbal sources, purified or modified version of phytochemicals and supplement.	No
	Mechanistic study (ie. bioavailability or mechanistic feeding study)	No
	Details for tool to estimated dietary polyphenol intake was not included	No
Access	Full-text article accessible	Yes

2.3.2 Data extraction and management

The studies were grouped, described and evaluated in accordance to their methodological similarities. Included studies were summarised in a tabular form, outlining study design, key feature of sample size and population, food intake assessed, reported polyphenol types and tools used for estimating dietary polyphenols intake.

2.3.3 Assessment of risk of bias in included studies

One review author (VG) assessed the quality for each study using the criteria outlined *in the Academy of Nutrition and Dietetics Evidence Analysis Manual 2012*, which critically appraises the quality of included studies. The checklist considers issues related to relevance and validity of included studies such as relevance improve current practice, randomisation, allocation concealment, blinding, intervention description, validity and reliability of measurements, missing data, selective reporting etc. When information in the studies was not sufficient, an attempt to contact the study authors was made to request further details. Studies were scored as positive, neutral or negative and were not excluded on the grounds of their quality.

3.0 Results and discussion:

A total of 2311 were identified from the searches conducted. A PRISMA flow chart of the search strategy and selection process was developed (Figure 2) which identified 157 studies to be included in the review. The full summary of these included studies can be found in the Data in Brief materials (Probst & Guan, 2016) for this manuscript.

3.1 Study characteristics and quality

Approximately 30% of studies were from case-control (n=44) and cross-sectional (n=48) study designs, respectively. The remainder of studies (n=65, 41%) were cohort studies. Included studies were from 24 different countries and 26% of studies from the United States (n=41) and approximately 13% of the studies from Japan. The majority of studies (n=130, 83%) used a food frequency questionnaire form of dietary assessment to estimate dietary polyphenols intake with 100% (n=44) of included case-control studies using this form of assessment. Food record (n=12) and 24-hour recall (n=8) dietary assessment methods also

were applied. Approximately 80% of studies (n=123) assessed intake in relation to the whole of diet rather than a single food item. Of the single food items specifically studied most were soy foods and legumes foods with few studies only focused on other key sources related to specific polyphenol subclasses eg. fruit, vegetables, tea, chocolate.

Upon assessing the quality of the published studies, there was one study was rated neutral using the Quality Criteria Checklist, due to the low response rate of dietary intake assessment in the cohort. This may imply the estimation of polyphenol intake was subject to bias. The remainder of studies were rated as positive. Additionally, the validation of selected dietary assessment tool was widely described in the studies.

3.2 Reported polyphenol subclasses

Figure 3 shows the distribution of reported polyphenols and subclasses from the studies. Isoflavonones were the most commonly reported polyphenol subclass. Approximately 35% (n=55) of studies reported total isoflavonones from the whole diet and 19% (n=30) of studies reported soy isoflavones intake. Approximately 80% (n=25) of studies focused on reported soy isoflavone intake were conducted in countries from Asia, while only 24% (n=13) of total isoflavones studies overall were from Asia. There were also a further 23% (n=36) of studies that reported isoflavone subclasses, genistein and daidzein. Approximately half of those (n=16) which reported isoflavone subclasses also reported their plant precursors, biochanin A and formononetin.

The second most common group was the flavonoid subclass. One quarter of studies (n=39) reported total flavonoid intake with a similar amount of studies also reporting individual

flavonoid compounds (n=42, 27%). Approximately 10% of these studies investigated multiple subclasses with seven (n=13), six (n=12) and five (n=14) subclasses, respectively identified. However, no reported pattern was revealed between reporting of individual flavonoids compounds and flavonoid subclasses.

Thirdly, one fifth of studies (n=33) provided intake information on total lignans, with half of these studies (n=18) reporting plant and/or mammalian lignans intake. Although flavonoids and lignans were widely reported, total polyphenols (n=6) and other subclasses of polyphenols (n=5) were rarely reported. Additionally, a total of only seven studies estimated dietary carotenoids intake despite some carotenoids appearing in traditional reference and survey food composition databases.

3.3 Tools used to estimate dietary polyphenols intake

Published literature was the most commonly identified tool used to translate food to polyphenol intake information. When considering the specific subclasses of polyphenols, identified tools for estimating dietary polyphenols and carotenoids intake included publicly accessible polyphenol (n=8) and carotenoid (n=1) databases, published database, published literature, and published analytical data based on local food items and analytical experiments. The identified tools, databases and food composition tables and their frequency of usage are presented as Table 2.

Approximately 60% (n=98) of studies included relied on publicly accessible databases or food composition tables to estimate dietary polyphenols intake. There were five studies identified that were using six food composition databases to assess intake. Approximately 60% (n=93) of studies applied only one tool, while 20% (n=31) of studies employed two and 13% (n=21) used three database or tool combinations.

Table 2: Identified tools and applied frequency for analysis of dietary intake of polyphenols

	Tools used	Frequency[a]	% of total (n=157)
1.	Published literature	53	34
2.	US Department of Agriculture- USDA Database for the Flavonoid Content of Selected Foods	43	27
3.	US Department of Agriculture- Iowa State University Database on the Isoflavone Content of Foods	30	19
4.	Published analytical data	28	18
5.	US Department of Agriculture-USDA Database for the Proanthocyanidin Content of Selected Foods	23	15
6.	The Phenol-Explorer database	17	11
7.	Published analytical data based on local food items	15	10
8.	China Food Composition Table	11	7
9.	Japan Food Composition Table	11	7
10.	UK Food Standards Agency Food Composition Database on phytoestrogens	7	4
11.	US Department of Agriculture-USDA database for the isoflavone content of selected foods	5	3

12.	Published database	4	3
13.	Analytical experiment	3	2
14.	Food Composition Tables Maintained from the University of Hawaii Cancer Center	3	2
15.	US Department of Agriculture-USDA national nutrient database for standard reference	3	2

[a] Frequency count includes studies where more than one tool was used to estimate intake.

When estimating soy isoflavone intake specifically, only one tool was used. The commonly used tools for isoflavone were the Chinese Food Composition Table, Japanese Food Composition Table, published literature or published analytical data based on local food items. Conversely, combinations of USDA databases were applied to assess isoflavone intake from the whole diet. In addition, plant precursors of isoflavone, genistein and daidzein were estimated using published literature or analytical data, rather than available databases.

When estimating seven subclasses of flavonoids at least three USDA databases or a combination of databases and published literature or analytical data were used. When studies assessing five or six subclasses of flavonoids, it was found that at least two USDA databases or a combination of databases and published literature or analytical data were applied. Either a combination of USDA databases or single published literature or published databases were more likely to be used to estimate dietary lignans intake. In addition, retention methods were reported to be the most commonly used method to expand the available food composition data of polyphenols to fit the reported food source (n=14).

While to the authors knowledge this review is the first to report on the food composition databases used in estimated intake studies, some earlier work has occurred in relation to databases available for phytochemicals (Scalbert et al., 2011). The previous review was particularly focused on the chemical structures, occurrence and concentrations in foods and also addressed metabolism in humans and animals and surrogate markers of health and focused its review on clinical trials which were not considered in the current review. There were synergies between two reviews though the work of Sclabert and colleagues did specify the particular components included in each of the food composition databases related generally to phytochemicals rather than their use in practice.

Both reviews are in agreement however with the need for flexible databases suited to the needs of the compounds. Where possible these databases should be able to be queried, contain a component of interactivity while maintaining the reliability and quality of the included components. For this to occur global efforts are required in relation to the terminology used, their application to practice and the suitability of particular data to regions which are geographically different such as for Australia.

4.0 Conclusion:

This review highlights the importance of publicly accessible food composition databases for estimation of dietary polyphenol intake. Despite the need for geographically specific data for these compounds, this review demonstrates that the USDA databases are most commonly applied despite the location of the study. There is a need for more geographically specific food composition databases at a global level with a consistent approach employed for their development. In parallel given the polyphenolic class of flavonoids, and its subsequent subclasses, are of particular interest to research examining various health outcomes, future studies could further highlight the methods of measurement pertaining to flavonoids intake, including biomarker data. This review also provides a systematic reference to the available tools to estimate dietary intake of polyphenols allowing researchers to determine the publicly available database which is most suited to the needs the study. This further demonstrates the need for researchers to disseminate their food composition data findings to improve accessibility to high quality data and reduce the privatisation of research outcomes.

Acknowledgements:

This work was supported by a National Health and Medical Research Council Translating Research into Practice Research Fellowship (APP1072484); NHMRC, Canberra Australia.

References:

Australian Bureau of Statistics. (2012). Daily intake of fruit and vegetables, 4338.0 - Profiles of Health, Australia, 2011-13S. Retrieved from

http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4338.0main+features202011-13

Australian Institute of Health and Welfare. (2013). AIHW analysis of the 2001 National Health Survey.

. Retrieved from http://www.aihw.gov.au/risk-factors-low-fruit-and-vegetable-consumption/

Beecher, G. R. (2003). Overview of dietary flavonoids: nomenclature, occurrence and intake. *Journal of Nutrition, 133*(10), 3248S-3254S.

Chun, O. K., Lee, S. G., Wang, Y., Vance, T., & Song, W. O. (2012). Estimated flavonoid intake of the elderly in the United States and around the world. *J Nutr Gerontol Geriatr, 31*(3), 190-205. doi:10.1080/21551197.2012.702530

Commenges, D., Scotet, V., Renaud, S., Jacqmin-Gadda, H., Barberger-Gateau, P., & Dartigues, J. F. (2000). Intake of flavonoids and risk of dementia. *European Journal of Epidemiology, 16*(4), 357-363.

Department of Health and Ageing, & National Health and Medical Research Council. (2011). A review of the evidence to address targeted questions to inform the revision of the Australian Dietary Guidelines: Evidence Statements Retrieved from http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/n55d_australian_dietary_guidelines_evidence_report.pdf

Dietitians Association of Australia. (2010). National Competency Standards for Entry Level Dietitians in Australia. Retrieved from http://daa.collaborative.net.au/files/WorkingandStudying/2010_ELC_Brochure_Inside_Word_VersionEXACT.pdf

Dwyer, J. T., & Peterson, J. J. (2002). Measuring flavonoid intake: need for advanced tools. *Public Health Nutrition, 5*(6a), 925-930. doi:doi:10.1079/PHN2002373

Hooper, L., Kroon, P. A., Rimm, E. B., Cohn, J. S., Harvey, I., Le Cornu, K. A., . . . Cassidy, A. (2008). Flavonoids, flavonoid-rich foods, and cardiovascular risk: a meta-analysis of randomized controlled trials. *American Journal of Clinical Nutrition, 88*(1), 38-50.

Knekt, P., Kumpulainen, J., Jarvinen, R., Rissanen, H., Heliovaara, M., Reunanen, A., . . . Aromaa, A. (2002). Flavonoid intake and risk of chronic diseases. *American Journal of Clinical Nutrition, 76*(3), 560-568.

Magarey, A., McKean, S., & Daniels, L. (2006). Evaluation of fruit and vegetable intakes of Australian adults: the National Nutrition Survey 1995. *Australian and New Zealand Journal of Public Health, 30*(1), 32-37.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement* (Vol. 339).

Monteiro, M. S., Carvalho, M., Bastos, M. L., & Guedes de Pinho, P. (2013). Metabolomics analysis for biomarker discovery: advances and challenges. *Current Medicinal Chemistry, 20*(2), 257-271.

O'Gorman, A., Gibbons, H., & Brennan, L. (2013). Metabolomics in the identification of biomarkers of dietary intake. *Computational and Structural Biotechnology Journal, 4*, e201301004. doi:10.5936/csbj.201301004

Park, E. J., & Pezzuto, J. M. (2012). Flavonoids in cancer prevention. *Anti-Cancer Agents in Medicinal Chemistry, 12*(8), 836-851.

Probst, Y., & Guan, V. (2016). Summary table of studies for dietary polyphenol intakes. *Food Chemistry, Submitted*(Data in brief).

Radtke, J., Linseisen, J., & Wolfram, G. (2002). Fasting plasma concentrations of selected flavonoids as markers of their ordinary dietary intake. *European Journal of Nutrition, 41*(5), 203-209. doi:10.1007/s00394-002-0377-z

Scalbert, A., Andres-Lacueva, C., Arita, M., Kroon, P., Manach, C., Urpi-Sarda, M., & Wishart, D. (2011). Databases on food phytochemicals and their health-promoting effects. *Journal of Agricultural and Food Chemistry, 59*(9), 4331-4348. doi:10.1021/jf200591d

Tapsell, L. C., Dunning, A., Warensjo, E., Lyons-Wall, P., & Dehlsen, K. (2014). Effects of vegetable consumption on weight loss: a review of the evidence with implications for design of randomized controlled trials. *Critical Reviews in Food Science and Nutrition, 54*(12), 1529-1538. doi:10.1080/10408398.2011.642029

Ulaszewska, M. M., Trost, K., Stanstrup, J., Tuohy, K. M., Franceschi, P., Chong, M. F.-F., . . . Mattivi, F. (2016). Urinary metabolomic profiling to identify biomarkers of a flavonoid-rich and flavonoid-poor fruits and vegetables diet in adults: the FLAVURS trial. *Metabolomics, 12*(2), 32. doi:10.1007/s11306-015-0935-z

Yokota, R. T., Miyazaki, E. S., & Ito, M. K. (2010). Applying the triads method in the validation of dietary intake using biomarkers. *Cadernos de Saude Publica, 26*(11), 2027-2037.

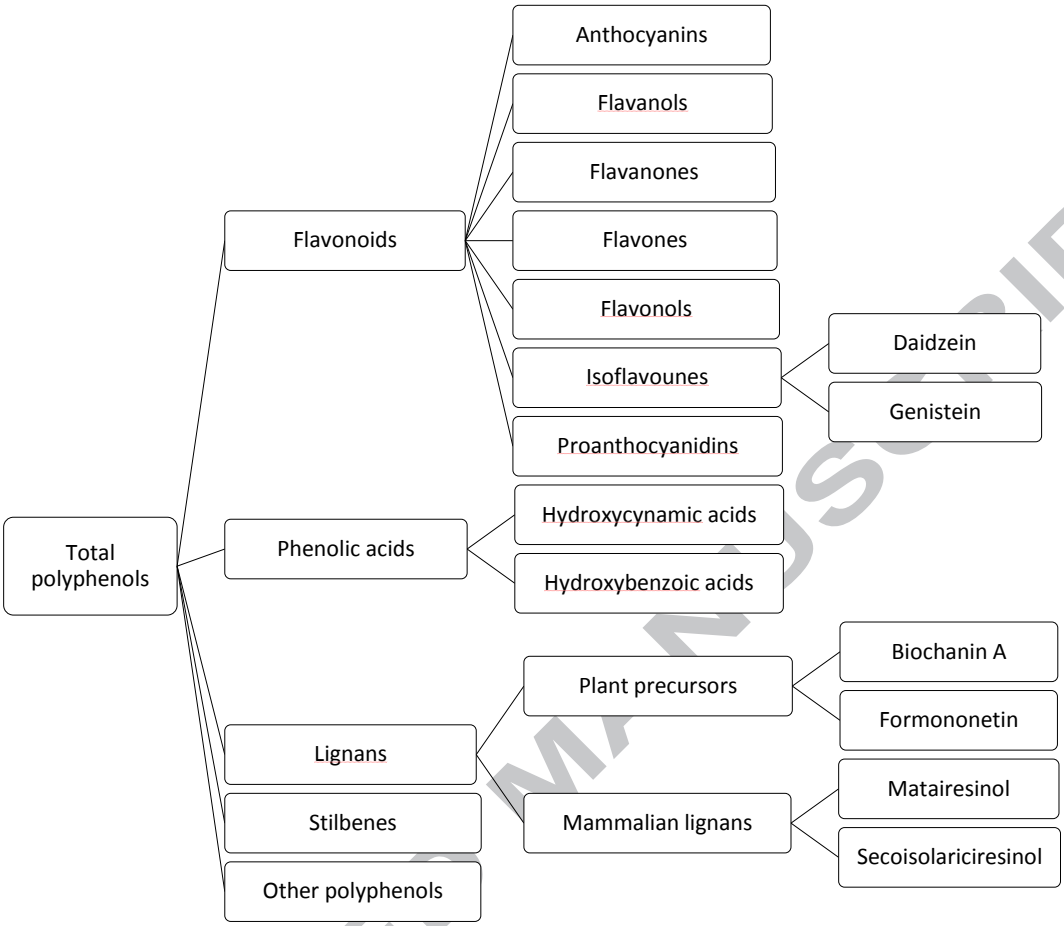
Figure captions:

Figure 1: Polyphenol subclasses

Figure 2: Representation of the number of studies per polyphenol class and sub-class

Figure 3: PRISMA flow diagram of the number of studies extracted for review

Figure(s) 1

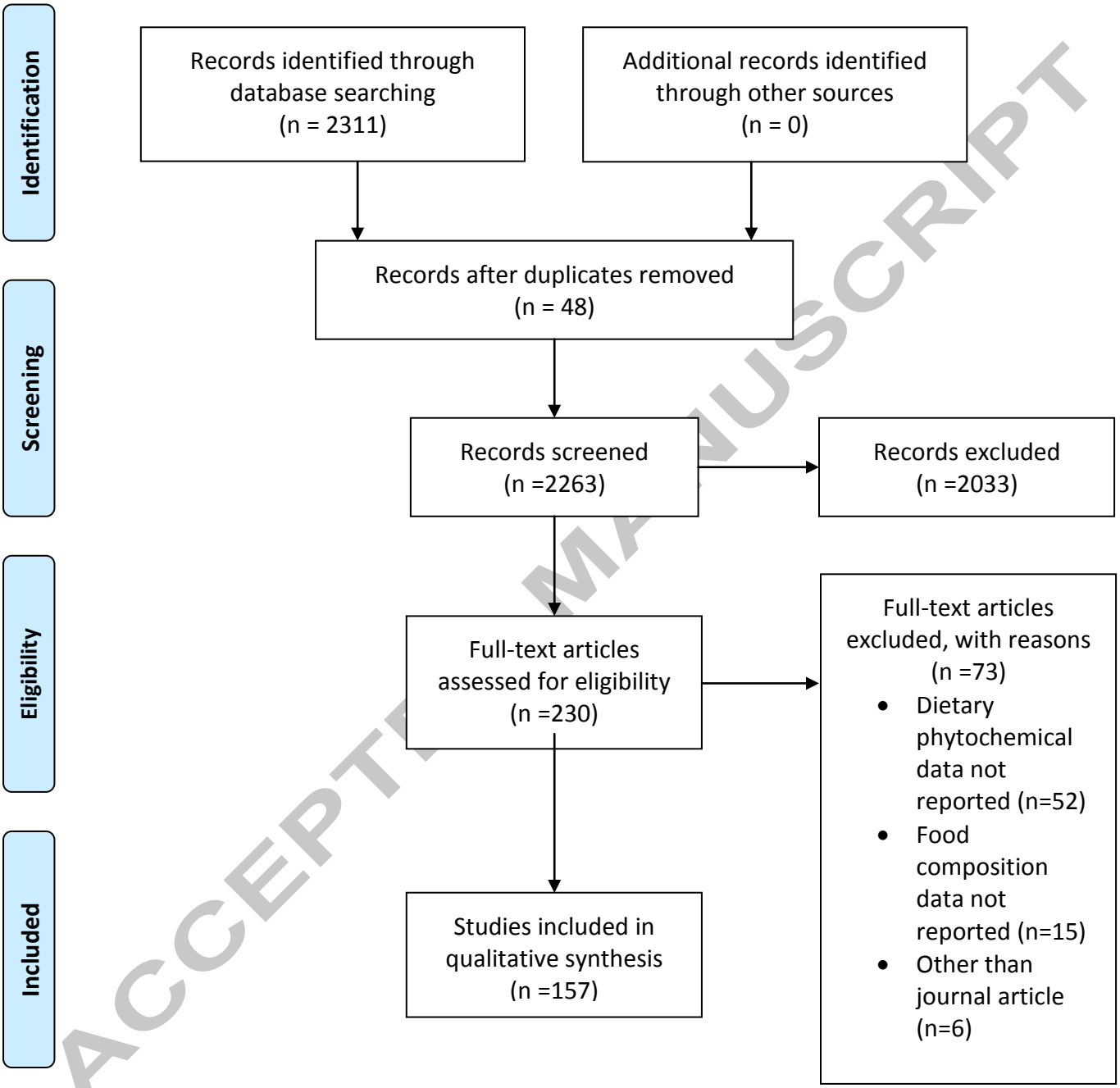


Figure(s) 2

Polyphenol class and subclass	Distribution of studies[a]
Flavonoids	
Total flavonoids	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●
Seven subclasses	●●●●●●●●●●●●●
Six subclasses	●●●●●●●●●●●●
Five subclasses	●●●●●●●●●●●●●●
Four subclasses	●●●
Three subclass	●●●●
Two subclasses	●●
Individual flavonoids compounds	●●
Isoflavones	
Total isoflavones	●●●
Soy isoflavones	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●
Subclass (genistein, daidzein)	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●
Plant precursors (biochanin A, formononetin)	●●●●●●●●●●●●●●●●
Lignans	
Total lignans	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●
Plant lignans (matairesinol, secoisolariciresinol) and/or mammalian lignans (enterolactone , lariciresinol, pinoresinol, syringaresinol, medioresinol, enterodiol, equol)	●●●●●●●●●●●●●●●●●●
Other polyphenols	
Total polyphenols	●●●●●●
Other polyphenols	●●●●●

[a] **Total number of included studies n=157, ●** = one study

Figure(s) 3



Highlights

- US food composition databases are most common for analysis of phytochemical intakes
- Case control studies use food frequency questionnaires for polyphenol analysis
- Total isoflavones, flavonoids and lignans are the most commonly studied subclasses
- Retention methods are used to expand existing food composition databases