A newly constructed and validated isoflavone database for the assessment of total genistein and daidzein intake

Margaret R. Ritchie1*, John H. Cummings2, Michael S. Morton3, C. Michael Steel1, Caroline Bolton-Smith4 and Andrew C. Riches1

1Bute Medical School, University of St Andrews, St Andrews, Fife KY16 9TS, UK
2Division of Neuroscience and Pharmacology, Ninewells Hospital, Dundee, UK
3Department of Medical Biochemistry, University of Cardiff College of Medicine, Cardiff, UK
414 Six Mile Bottom Road, West Wratting, Cambridge, UK

(Received 23 January 2005 – Revised 16 June 2005 – Accepted 16 August 2005)

The principal phyto-oestrogens (PO) in food are isoflavones, lignans, coumestans and prenylated flavonoids, with isoflavones and lignans being the most commonly found in UK diets. Until recently obtaining accurate data on the PO content of foods was hampered by lack of suitable analytical methods and validation techniques. Furthermore, although PO data exist for some foods, these may not be available in the UK. The aim of the present study was to construct a new, comprehensive isoflavone (total genistein + daidzein) database. Using data, mainly from recent GC–MS analysis, for approximately 300 foods available in the UK, and extensive recipe calculations, a new database was constructed containing approximately 6000 foods allocated an isoflavone value. By analysing 7 d weighed food diaries, the database was subsequently used to estimate isoflavone intake in two groups of healthy volunteers, omnivores (n 9) and vegetarians (n 10). Mean isoflavone intake in the vegetarian and omnivorous group was 7·4 (SEM 3·05) and 1·2 (SEM 0·43) mg/d, respectively. Mean intake for the total group was 4·5 (SEM 1·89) mg/d. Main food sources of isoflavones for the vegetarian group were soya milk (plain), meat-substitute foods containing textured vegetable protein and soya protein isolate, soya sauce, raisins and scones. Main food sources of isoflavones for the omnivorous group were soya yogurts, wholemeal bread and rolls, white bread and rolls, garlic bread, nan bread and brown bread, sultanas and scones.

Isoflavone database: Recipe calculations: Dietary intake

There is increasing interest in phytochemicals and their role in disease prevention. Phytochemicals include dietary phyto-oestrogens (PO) such as the isoflavones genistein and daidzein, present in fruits, nuts, peas and beans, lentils, chickpeas and soya (Setchell et al. 1987; Franke et al. 1995; Mazur et al. 1996; Reinli & Block, 1996; Liggins et al. 1998, 2000a,b, 2002; Mazur & Adlercreutz, 1998), lignans, which are present in cereals, beans, peas, tea (black and green), wine and strawberries (Thompson et al. 1991; Obermeyer et al. 1995; Adlercreutz & Mazur, 1997; Nesbitt & Thompson, 1997; Bingham et al. 1998; Mazur & Adlercreutz, 1998, Mazur et al. 1998a,b, 2000; Glitso et al. 2000) and prenylated flavonoids found in hops (Rong et al. 2000).

Soya is recognised as the major dietary source of PO (Bingham et al. 1998; Mazur & Adlercreutz, 1998, 2000; United States Department of Agriculture & Agricultural Research Service, 2004) and soya-based products have been shown to contain significant quantities of the isoflavones genistein and daidzein (Murphy et al. 1999; Pillow et al. 1999). Since genistein and daidzein are the most prevalent dietary isoflavones, values for their amounts in foods were used in the construction of the new isoflavone database.

There are a number of factors which can affect the isoflavone concentration in plants such as plant species and strain, crop year and geographical location (Franke et al. 1995). Processing can also affect isoflavone concentrations (Wang et al. 1990).

Although several databases have been constructed containing the isoflavone content of foods (Pillow et al. 1999; Horn-Ross et al. 2000; United States Department of Agriculture & Agricultural Research Service, 2004; Vegetable Estrogens in Nutrition and Skeleton (VENUS), 2005), there is no current isoflavone database comprehensive enough to be used in epidemiological studies in the UK which assess the dietary impact of isoflavones on the health of a population.

The new isoflavone database has been constructed using a combination of techniques including results from previous analyses (Knight et al. 1998; Mazur & Adlercreutz 1998; Pillow et al. 1999; Horn-Ross et al. 2000), more recent analyses (Liggins et al. 1998, 2000a,b, 2002), recipe calculations and

Abbreviations: EI, energy intake; FC, food code; M&W, McCance & Widdowson; PO, phyto-oestrogen; VENUS, Vegetable Estrogens in Nutrition and Skeleton.

* Corresponding author: Dr Margaret R. Ritchie, fax +44 1334 463482, email mrr3@st-andrews.ac.uk
nutrient information obtained from several major supermarket chains in Tayside, UK, including Sainsburys, Asda, Tesco, Morrisons, Iceland and Farmfoods. Foods from supermarkets were sampled and ingredient lists and variations in ingredient lists were checked by telephone calls to supermarket head offices and stores in other areas within the UK. Nutrient information was also obtained from UK food manufacturers relating to over 1600 bakery and frozen products.

Methods
Initially the sum of genistein and daidzein values were allocated to a range of foods on the database, many of which were identified from over 1000 food diaries, 24 h dietary recalls and food-frequency questionnaires obtained during several studies over a 4-year period (Ritchie et al. 2004a,b,c; Heald et al. 2005).

A database should be constructed with, as far as possible, one method of analysis being used throughout (Ziegler, 2001) and the analysis should be recent, reliable and sensitive (Ovaskainen et al. 1996; Ziegler, 2001). For these reasons, values of genistein and daidzein used in construction of the UK isoflavone database were mainly obtained from recent analyses carried out by Liggins et al. (1998, 2000a,b, 2002) using GC–MS. Although some HPLC values for the isoflavone content of foods were used in the database construction, HPLC analysis of foods was not considered sensitive, reliable and accurate enough to be used with great confidence. For example, in Table 1, the results for analyses performed by researchers (Adlercreutz & Mazur, 1997; Mazur & Adlercreutz, 1998; Liggins et al. 2002) largely concur; however, those for green split peas and mung bean sprouts differ considerably from values published by Franke et al. (1995).

The database also included brand names and recipe dishes (Ovaskainen et al. 1996).

### Construction of new isoflavone database

Since genistein and daidzein are the principal isoflavones in the diet (Pillow et al. 1999), comprising more than 90% of the intake of oestrogenic isoflavones, the sum of levels of these compounds in foods was allocated to corresponding foods in the database. In order to distinguish values to be used for the database construction, all the isoflavone data gathered from research papers underwent quality assurance (Ovaskainen et al. 1996; Ziegler, 2001) involving a number of steps.

For foods analysed, did the author specify the following?
- type of food and country of origin;
- McCance & Widdowson (M&W) code;
- if food was raw or cooked;
- if the results were for dry weight or wet weight (as is);
- the type of internal standard used;
- method of analysis (including CV quoted, multiple sample analyses for accuracy, adequate level of detection).

It was essential that all these criteria were fulfilled in order for results to be considered for inclusion in the database (exceptions were ‘new user foods’ without M&W codes).

Table 2 is an example of the quality assurance carried out. Comparison of genistein and daidzein concentrations in cereals by different methods of analysis was also performed and values are presented in Table 3.

### Table 1. Comparison of genistein and daidzein values obtained by different studies for selected vegetables (adapted from Liggins et al. 2000b)

<table>
<thead>
<tr>
<th>Food</th>
<th>Study</th>
<th>Daidzein (μg per 100 g)</th>
<th>Genistein (μg per 100 g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soyabees boiled*</td>
<td>Liggins et al. (2002)</td>
<td>15 000</td>
<td>32 000</td>
</tr>
<tr>
<td>Soyabees boiled*</td>
<td>Mazur et al. (1998a)</td>
<td>20 000</td>
<td>32 000</td>
</tr>
<tr>
<td>Miso</td>
<td>Liggins et al. (2002)</td>
<td>59 000</td>
<td>67 000</td>
</tr>
<tr>
<td>Bean sprouts (mung)*</td>
<td>Liggins et al. (2002)</td>
<td>39 000</td>
<td>68 000</td>
</tr>
<tr>
<td>Bean sprouts (mung)†</td>
<td>Adlercreutz &amp; Mazur (1997)</td>
<td>800</td>
<td>1900</td>
</tr>
<tr>
<td>Bean sprouts (mung)†</td>
<td>Franke et al. (1995)</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Broad beans*</td>
<td>Liggins et al. (2002)</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Broad beans†</td>
<td>Mazur &amp; Adlercreutz (1998)</td>
<td>24</td>
<td>tr</td>
</tr>
<tr>
<td>Broccoli, calabrese*</td>
<td>Liggins et al. (2002)</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Broccoli, calabrese†</td>
<td>Adlercreutz &amp; Mazur (1997)</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Carrots†</td>
<td>Liggins et al. (2002)</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Carrots†</td>
<td>Adlercreutz &amp; Mazur (1997)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cranberry*</td>
<td>Liggins et al. (2002)</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>Cranberry†</td>
<td>Adlercreutz &amp; Mazur (1997)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Green split peas*</td>
<td>Liggins et al. (2002)</td>
<td>13</td>
<td>35</td>
</tr>
<tr>
<td>Green split peas†</td>
<td>Mazur et al. (1998a)</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Green split peas†</td>
<td>Franke et al. (1995)</td>
<td>7 300</td>
<td>nd</td>
</tr>
<tr>
<td>Mushrooms*</td>
<td>Liggins et al. (2002)</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Mushrooms†</td>
<td>Adlercreutz &amp; Mazur (1997)</td>
<td>20</td>
<td>12</td>
</tr>
</tbody>
</table>

nd, not detected; tr, isoflavone identified but could not be quantified.
* Concentrations expressed on a wet-weight basis.
† Concentrations expressed on a dry-weight basis.

Recipe calculations

All isoflavone values quoted represent the sum of genistein + daidzein and are in μg per 100 g for wet weight.

The same mathematical formula was used for calculating isoflavone content of bakery products involving the ratio of protein content of the required bakery product. For example, brown rolls crusty (A)/protein content of the same bread...
type, for example, brown bread (B) × isoflavone concentration of bread type, as measured by Liggins et al. (2000a, b, 2002). It was assumed that isoflavone content in bakery products is proportional to protein content. See Table 4 and Appendix 1.

Calculations (Appendix 1) for estimation of isoflavone content of composite dishes were carried out in a similar manner to those used for beef sausages (Table 5), fish pie and carrot and bean salad dishes.

Three brand-name soya milks were analysed for total genistein + daidzein content (by M. S. M.) using GC–MS. These foods were entered onto the database as ‘new user’ foods (see Table 6).

The isoflavone content of sweetened soya milk (M&W food code (FC) 12 043) $= (A + B)/2 \mu g/100 g$;

So, the estimated isoflavone content of sweetened soya milk $= 8155 \mu g/100 g$.

The isoflavone content of plain soya milk was calculated by averaging isoflavone values for soya milk (published and unpublished values). The values used are listed in Table 7.

Isoflavone (genistein + daidzein) estimated values for foods

These were obtained in two ways:

- by calculation using a recipe;
- by comparison with isoflavone values for similar foods which had been analysed.

For example, the isoflavone content of raisins (M&W FC 14 242) $= 183.6 \mu g/100 g$ (A);

The isoflavone content of currants (M&W FC 14 074) $= 224.5 \mu g/100 g$ (B);

So, the estimated isoflavone content of sultanas $= (A + B)/2 = 204.1 \mu g/100 g$;

The isoflavone content of database addition (M&W FC 14 263) $= 204.1 \mu g/100 g$.

Further examples of estimated isoflavone values and corresponding foods used in allocation of isoflavone levels are listed in Table 8.

Isoflavone (genistein + daidzein) for ‘second generation’ soya-containing foods

Several foods containing soya extract which were consumed by volunteers did not appear on the database (i.e. no M&W FC) and subsequently had to be added. Furthermore, there was no reference to their isoflavone content in the literature. Isoflavone content of the food item was estimated using nutrient information present on food wrappers detailing the percentage composition. A ‘new user food’ code was also designated. An example of this is given in Appendix 2.

Assessment of dietary isoflavone intake using the database: rationale

Once constructed, it was necessary to test the database. The validity of the database was tested using the method of duplicate diet analysis (Ritchie et al. 2004a). The accuracy of the database was tested by using it to assess dietary isoflavone intake in healthy volunteers consuming a range of intakes (vegetarians and omnivores), i.e. two groups, with different anticipated isoflavone intakes.
The aim of the present study was to determine usual isoflavone intake in healthy vegetarians and omnivores living in the Dundee area using a 7 d weighed food diary.

<table>
<thead>
<tr>
<th>Food</th>
<th>Study</th>
<th>Daidzein (µg per 100 g)</th>
<th>Genistein (µg per 100 g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bread, white</td>
<td>Liggins et al. (2002)</td>
<td>135·6</td>
<td>157·2</td>
</tr>
<tr>
<td>Bread, white</td>
<td>Horn-Ross et al. (2000)</td>
<td>606·0</td>
<td></td>
</tr>
<tr>
<td>Bread, wholegrain</td>
<td>Liggins et al. (2002)</td>
<td>373·1</td>
<td>456·7</td>
</tr>
<tr>
<td>Bread, wholegrain</td>
<td>Horn-Ross et al. (2000)</td>
<td>155·8</td>
<td>141·8</td>
</tr>
<tr>
<td>Oatmeal</td>
<td>Liggins et al. (2002)</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>Oatmeal</td>
<td>Horn-Ross et al. (2000)</td>
<td>0 or trace</td>
<td>0 or trace</td>
</tr>
<tr>
<td>Oatmeal</td>
<td>Adlercreutz &amp; Mazur (1997)</td>
<td>3·5</td>
<td>6·0</td>
</tr>
</tbody>
</table>

nd, not detected.

<table>
<thead>
<tr>
<th>Food</th>
<th>M&amp;W food code</th>
<th>Protein (%)</th>
<th>Total G + D content (µg per100g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White rolls, soft</td>
<td>11124</td>
<td>9·2</td>
<td>320·7</td>
</tr>
<tr>
<td>Brown rolls, crusty</td>
<td>11118</td>
<td>10·3</td>
<td>635·7</td>
</tr>
<tr>
<td>Brown rolls, soft</td>
<td>11119</td>
<td>10·0</td>
<td>624·5</td>
</tr>
<tr>
<td>Wholemeal, rolls</td>
<td>11125</td>
<td>9·0</td>
<td>811·8</td>
</tr>
<tr>
<td>Croissants</td>
<td>11120</td>
<td>8·3</td>
<td>289·3</td>
</tr>
<tr>
<td>Chapatis made with fat</td>
<td>R11074</td>
<td>8·1</td>
<td>282·3</td>
</tr>
<tr>
<td>Chapatis made minus fat</td>
<td>46</td>
<td>7·3</td>
<td>254·5</td>
</tr>
<tr>
<td>Nan bread</td>
<td>11086</td>
<td>8·9</td>
<td>310·2</td>
</tr>
<tr>
<td>Pitta bread</td>
<td>11090</td>
<td>9·2</td>
<td>310·7</td>
</tr>
<tr>
<td>Rye bread*</td>
<td>48</td>
<td>–</td>
<td>33</td>
</tr>
</tbody>
</table>

M&W, McCance & Widdowson.  
* Assumes isoflavone value same as that for granary bread (Liggins et al. 2002).

<table>
<thead>
<tr>
<th>Food</th>
<th>M&amp;W food code</th>
<th>TVP (%)</th>
<th>Potato (%)</th>
<th>Total G + D content (µg per 100g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beef sausages, ‘Premium’</td>
<td>19095</td>
<td>0·85</td>
<td></td>
<td>580</td>
</tr>
<tr>
<td>Frankfurters</td>
<td>19100</td>
<td>0·99</td>
<td></td>
<td>676</td>
</tr>
<tr>
<td>Beefburgers, fried</td>
<td>19029</td>
<td>0·75</td>
<td></td>
<td>512</td>
</tr>
<tr>
<td>Beefburgers, low-fat</td>
<td>19037</td>
<td>0·75</td>
<td></td>
<td>512</td>
</tr>
<tr>
<td>Bridie or scotch pie</td>
<td>19053</td>
<td>0·98</td>
<td></td>
<td>253*</td>
</tr>
<tr>
<td>Cornish pastie</td>
<td>19056</td>
<td>0·98</td>
<td></td>
<td>253*</td>
</tr>
<tr>
<td>Fishcakes</td>
<td>16281</td>
<td>7·2</td>
<td></td>
<td>0·2</td>
</tr>
</tbody>
</table>

M&W, McCance & Widdowson; TVP, textured vegetable protein.  
* Assume G + D are found in meat only, which is 40 % of food.

Isoflavone intake assessed using new database

Table 4. Isoflavone (genistein (G) plus daidzein (D)) values for other bakery products

Table 5. Genistein (G) plus daidzein (D) values for other meat products

Table 6. Phyto-oestrogen (genistein and daidzein) values for branded soya milks (MS Morton, unpublished results)*

Study design and methods

Nineteen healthy volunteers (age range 19–76 years), seventeen female and two male were recruited by advertisement (poster) in Ninewells Hospital and Dundee University. The group consisted of nine omnivores, eight females, one male,
mean age 40 (range 19–76) years and ten vegetarians, nine females, one male; mean age 36 (range 21–56) years. Ethical approval was granted for the study and signed consent was obtained from each subject.

Subject weight and height were recorded at the start and end of the 1 week of weighed food recording (Bingham et al. 1988). BMI was calculated for each subject and BMR was calculated using Schofield equations (Schofield et al. 1985). Energy intake (EI) from the food diaries (using ‘Microdiet’; Downlee Systems Ltd., Downlee Lodge SK23 9UB, UK) was used to calculate the EI:BMR ratio for each volunteer. EI:BMR had to be greater than or equal to 1.2 to indicate subject compliance. Dietary composition was calculated using ‘Microdiet’ and individual and mean group isoflavone intakes were estimated using Excel (Microsoft Corporation, Redwood, WA, USA) and SPSS (SPSS Inc., Chicago, IL, USA) statistical analysis programs.

Isoflavone intake in the vegetarian and the omnivorous groups were recorded (Table 9) and compared by Student’s t test. The percentage contribution for each of the major PO-containing foods to mean intake of each group, i.e. vegetarian or omnivore, was calculated and foods were listed in order of greatest dietary isoflavone contributor to lowest contributor (Table 10).

**Results**

A database containing isoflavone (total genistein + daidzein) concentrations in μg per 100 g allocated to approximately 6000 foods was constructed and placed on the website at http://medicine.st-and.ac.uk/research/docs/ritchie/. To access the database the user name and password are gentian and violet respectively.

Where a value has been derived from a database containing genistein and daidzein values obtained by a variety of methods (such as in the United States Department of Agriculture database), the reference reflects this. Hence, if HPLC and GC–MS were used, this is represented as GC–MS/HPLC. Food reference codes are based on McCance & Widdowson’s The Composition of Foods, 4th ed. (Paul & Southgate, 1978) and 5th ed. (Royal Society of Chemistry, 1991), together with the 1980 and 1985 supplements.

The following symbols are used:

- N = currently no known value (from recent analysis)
- c = estimated value (usually from a similar food where genistein + daidzein concentrations have been measured)
- r = recipe used
- cr = value calculated using a recipe
- m = measured value
- T = source of value and rationale for using it are outlined (Ritchie, 2003)

**Subject compliance**

Mean subject weight (kg) at the start and end of the recording period was 67.1 (range 46.5–117) and 67.1 (range 46.6–116.0), which indicated good compliance regarding usual food intake.

Individual isoflavone intakes are represented graphically in Fig. 1.

### Table 7. Isoflavone (total genistein (G) plus daidzein (D)) values used for estimating the content of plain soya milk*

<table>
<thead>
<tr>
<th>Author</th>
<th>G + D content (μg per 100 g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morton et al. (2002)</td>
<td>17 840</td>
</tr>
<tr>
<td>unpublished results</td>
<td></td>
</tr>
<tr>
<td>Pumford et al. (2002)</td>
<td>14 600</td>
</tr>
<tr>
<td>Wiseman et al. (2002)</td>
<td>10 800</td>
</tr>
<tr>
<td>Wiseman et al. (2002)</td>
<td>12 600</td>
</tr>
<tr>
<td>Horn-Ross et al. (2000)</td>
<td>11 322</td>
</tr>
<tr>
<td>USDA–IUS database†</td>
<td>4710†</td>
</tr>
<tr>
<td>USDA–IUS database‡</td>
<td>10 501</td>
</tr>
</tbody>
</table>

USDA, United States Department of Agriculture; ISU, Iowa State University.

* Plain soya milk; McCance & Widdowson food code 12 042; mean phyto-oestrogen content 11 800 μg per 100 g.

† Iced soya milk.


### Table 8. Estimated genistein (G) plus daidzein (D) values for foods not analysed (FN) compared with similar foods which have been analysed (FA) (Liggins et al. 1998)

<table>
<thead>
<tr>
<th>FA for G + D</th>
<th>Measured G + D content (μg per 100 g)</th>
<th>FN for G + D</th>
<th>Estimated G + D content (μg per 100 g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readybrek</td>
<td>0</td>
<td>Porridge with milk</td>
<td>0</td>
</tr>
<tr>
<td>Chickpeas, cooked</td>
<td>24-1</td>
<td>Chickpeas, tinned</td>
<td>24-1</td>
</tr>
<tr>
<td>Lentils red, split</td>
<td>0-5</td>
<td>Lentils, green and brown</td>
<td>0-5</td>
</tr>
<tr>
<td>Carrots, canned</td>
<td>0</td>
<td>Carrots, new, boiled</td>
<td>0</td>
</tr>
<tr>
<td>Broccoli, purple</td>
<td>1-8</td>
<td>Broccoli, green</td>
<td>1-8</td>
</tr>
<tr>
<td>Tomatoes, raw</td>
<td>3-25</td>
<td>Tomatoes, cherry</td>
<td>3-25</td>
</tr>
<tr>
<td>Peppers, green</td>
<td>0</td>
<td>Peppers, red</td>
<td>0</td>
</tr>
<tr>
<td>Fruit cocktail in syrup</td>
<td>1-54</td>
<td>Fruit cocktail in juice</td>
<td>1-54</td>
</tr>
<tr>
<td>Oranges</td>
<td>0</td>
<td>Grapefruit</td>
<td>0</td>
</tr>
</tbody>
</table>
Main sources of dietary isoflavones and intake (mg/d) for each group

The main food sources of isoflavones and subject intakes from these sources (mg/d over the 7d recording) for the vegetarian group were soya milk (plain) (24.5–15.3), soya mince (3.8), meat-substitute foods containing textured vegetable protein and soya protein isolate (0.4), wholemeal bread and rolls (0.3), white bread and rolls (0.4), bread varieties such as croissants (0.2) and pitta breads (<0.1), beans (<0.1), raisins (<0.1) and soya sauce (<0.01).

The main food sources of isoflavones and subject intakes from these sources (mg/d over the 7d recording) for the omnivorous group were soya yogurts (2.8), wholemeal bread (0.3) and rolls (0.2), white bread (0.3) and rolls (0.2), varieties of bread such as garlic bread (0.1), nan bread (<0.1) and brown bread (0.2), sultanas (<0.1) and scones (<0.1).

Discussion

Assessment of dietary intake of specific nutrients or non-nutrients tends to be based on information obtained from food diaries, dietary recall or food-frequency questionnaires, all of which have to be analysed using a database containing appropriate values for foods consumed. Although dietary intake of isoflavones is correlated with intake of soya and soya-based foods (Verkasalo et al. 2001), a number of studies report apparently low dietary isoflavone intakes in Western diets (Jones et al. 1989; Verkasalo et al. 2001) due to the lack of genistein plus daidzein values for non-soya-based foods or the restricted range of non-soya-based foods included in the dietary assessment (Kirk et al. 1999). Examples of non-soya-based foods that contain genistein and daidzein are processed foods, bakery products and composite foods such as salads (containing raisins, rice, chickpeas, haricot beans, butter beans and beansprouts), desserts and mixed vegetable dishes.

The analysis required, of the extensive range of foods consumed in the UK, in order to create a fully comprehensive isoflavone database for use in epidemiological studies would not be feasible in terms of time and costs.

The current database was originally developed for the analysis of food-frequency questionnaires used in a study...
investigating isoflavone and prostate cancer risk (Food Standards Agency project number FS2069). It was assumed that this more comprehensive database would provide more accurate information about actual isoflavone intake across a range of intakes and, in particular, in populations not consuming a soya-based diet.

Other methods for assessing total isoflavone intake per d were employed by Clarke and colleagues (Clarke et al. 2003, 2004; Clarke & Lloyd, 2004). The researchers analysed vegetarian duplicate diets and food groups from the 1998 Total Diet Survey for isoflavones. There was considerable agreement between the two methods (isoflavone database and food group analysis) used for assessing the contribution on non-soya-based food groups to isoflavone intake. Clarke et al. (2004) report the isoflavone contribution of each food group to isoflavone intake from highest to lowest as bread, processed meat, fish products, cereals, fruit products and nuts, which is largely in agreement with the order indicated in the new database.

Comparison with other databases

The isoflavone databases produced by the United States Department of Agriculture and Iowa State University (United States Department of Agriculture & Agricultural Research Service, 2004) and Pillow et al. (1999) contain foods eaten in America but not obtainable in the UK, for example, ‘Bacon bits’, ‘Baco’s’, ‘Arrowhead Mills multigrain corn bread’, etc. Furthermore the isoflavone content of breads (white bread and wholegrain) in the database produced by Horn-Ross et al. (2000a) differ considerably from the values obtained by Liggins et al. (2002). This is probably because soya flour is added to bread (Liggins et al. 2002) and different countries use different amounts in their bakery products. For this reason, isoflavone values assigned by Horn-Ross et al. (2000a) to bakery products were not used in the database for foods eaten in the UK. The VENUS database contains isoflavone values for foods analysed using a variety of methods, internal standards and methods of detection. The lack of consistency in methods of analysis and restricted number of foods limit the use of this database in epidemiological studies. The advantages and strengths of the newly constructed database are due to the large number of UK foods which have been used in its construction. Many PO values are obtained from the GC–MS analysis, which is both reliable and recent. In addition, many values were derived from recipe calculations relating to the extensive range of foods sampled during the database construction. Furthermore, the database includes a range of bakery products such as toast, muffins, croissants, Chelsea buns, pitta bread and nan bread, a range of salads such as bean salad, Florida salad, potato salad, and coleslaw and a range of combined dishes such as fruit pie and wholemeal pancakes stuffed with vegetables. As such it is the first comprehensive isoflavone database to be constructed for use in UK-based studies assessing isoflavone intake.

Dietary intake of isoflavones

Analysis of food diaries using the database estimated a dietary isoflavone intake of 1·2 (range 0·2–3·5) mg/d for the omnivorous group and a mean dietary intake for the total group of 4·45 mg/d. This is higher than the value of <1 mg/d quoted by Jones et al. (1989), as a result of HPLC analysis of diet samples obtained during the 1987 British Total Diet Survey.

Since 1989, the addition of soya products to food items in supermarkets has increased and it is probable that the Total Diet Survey underestimates the mean intake of isoflavones in the UK. Horn-Ross et al. (2000) assessed the mean isoflavone intake in 447 non-Asian women (age 50–79 years) in the San Francisco Bay area, USA, using a modified food-frequency questionnaire and a newly constructed database. Mean intakes of genistein and daidzein were estimated as 1·5 and 1·3 mg/d, respectively, resulting in a total isoflavone intake of genistein plus daidzein equal to 2·8 mg/d.

This value is similar to that of 3 mg/d proposed by Clarke et al. (2003) based on analysis of samples obtained from the 1998 UK Total Diet Survey. This intake was obtained exclusively from bread and processed meats. It is also in agreement with 3·5 mg as being the upper range of low isoflavone intakes reported in the present study. During validation of the database Ritchie et al. (2004a) indicated that, at low isoflavone intakes, the database underestimates intake by 1·2–2·2 mg. By adding this to the mean intake of 1·2 mg/d, estimated using the database, isoflavone intake at low intakes (mainly due to breads) agrees with the value of 3 mg quoted by Clarke et al. (2003).

Using the newly constructed database, a mean vegetarian isoflavone intake of 7·4 mg/d was estimated. This was expected since vegetarian diets were assumed to contain more plant-based foods and possibly more food items consisting of meat substitutes and likely to contain soya.

Mean daily intakes of 10·5 mg/d were measured in thirty-five duplicate vegetarian diets by Clarke et al. (2003) using LC-MS. Reasons for the difference between the estimated mean isoflavone intake for the vegetarian group in the present study and the measured intake may include the small sample size in the present study. In addition, one vegetarian in the present study did not eat bread and wheat products. These foods are major sources of PO in non-Asian diets (Horn-Ross et al. 2000).

The main food sources of isoflavones for each group in the present study were soya milks and yogurts, soya and textured vegetable protein-based foods, breads, dried fruit. This is in agreement with work carried out by Horn-Ross et al. (2000b) who found that tofu, doughnuts, soya milk and bread were primary sources of PO for post-menopausal women in the USA.

The standard errors of the mean associated with estimated isoflavone intake were high (Table 9) due to the small sample size and large range of estimated intakes. However, the aim of the present study was to construct a new, more comprehensive isoflavone database and test it, both to find out the number of foods requiring isoflavone values and how the estimated intakes compared with published values. For the present study 5570 foods on the database were initially assigned an isoflavone value, of which 4887 occur as distinct foods. A total of 4852 foods are allocated a value of zero isoflavone due to estimates, measured values or isoflavone content unknown.

By adding estimated values, the final database contains 5951 foods assigned an isoflavone value.
There is good correlation (r 0.98) between estimated isoflavone intake using the database and measured isoflavone intake from duplicate diet analysis (Ritchie et al. 2004a). Hence, results of duplicate diet analysis demonstrate the ability of the database to assess accurately dietary intakes of isoflavones.

This new database makes a valuable contribution to the requirement for a comprehensive isoflavone database available for use in epidemiological studies assessing the effect of isoflavone intake on health. As more information on the isoflavone content of foods becomes available, the database can be updated and added to. The construction of the new database and its availability, however, provides the initial step for future studies involving a more accurate assessment of dietary intakes of phytochemicals and their effects on health and disease prevention.

Acknowledgements
The Anonymous Trust, University of Dundee, supported the study.

References
Ritchie MR, Morton MS, Deighton N, et al. (2004b) Plasma and urinary concentrations of isoflavones as biomarkers of phyto-oestrogen intake assessed using new database 211.


Appendix 1. Isoflavone (genistein plus daidzein) content for bakery products

**White bread, toasted; food code 11106**

Genistein plus daidzein (G + D) content of white bread (WB) is 292·8 μg per 100 g.

So, G + D content of WB, toasted = (% protein WB, toasted × 292·8)/% protein WB

\[
\frac{9·3 \times 292·8}{8·4} = 352·2 \mu g \text{ per 100 g.}
\]

**Brown bread, toasted; food code 11073**

G + D content of brown bread (BB) is 524·6 μg per 100 g.

So, G + D content of BB, toasted = (% protein BB, toasted × 524·6)/% protein BB

\[
\frac{10·4 \times 524·6}{8·5} = 641·9 \mu g \text{ per 100 g.}
\]

**Wholemeal bread, toasted; food code 11117**

G + D content of wholemeal bread (WHB) is 829·8 μg per 100 g.

So, G + D content of WHB, toasted = (% protein WHB, toasted × 829·8)/% protein WHB

\[
\frac{10·8 \times 829·8}{9·2} = 974·1 \mu g \text{ per 100 g.}
\]

**White rolls, crusty; food code 11123**

G + D content of white rolls, crusty (WRC) is 292·8 μg per 100 g (assume same as WB).

So, G + D content of WRC = (% protein WRC × 292·8)/% protein WB

\[
\frac{9·3 \times 292·8}{8·4} = 379·9 \mu g \text{ per 100 g.}
\]

**Appendix 2. Isoflavone (genistein plus daidzein) values for meat products**

**Pork sausages (brand name); food code 19091**

Pork and beef sausages ‘Economy’ contain 0·85 % soya concentrate.

Assume that the G + D content of soya concentrate is the G + D content of textured vegetable protein (TVP).

G + D content of TVP = \( \frac{(68 600 + 68 000)}{2} = 68 300 \mu g \) per 100 g.

In 100 g sausages, weight of TVP = 0·85 g.

G + D content of 100 g TVP = 68 300 μg.

So, in 0·85 g of TVP we have 0·85/100\% G + D.

G + D content of 100 g sausages = 580 μg per 100 g.

The same formula was used for beef sausages ‘Premium’, beef olives, beefburgers, bridie or scotch pie, and Lorne sausage. The foods and associated G + D values are listed in Table 5.

**Appendix 3. Isoflavone (genistein plus daidzein) values for fish products**

**Recipe for fish pie; McCance & Widdowson food code 16294**

200 g cooked cod

150 ml milk

400 g mashed potato

15 g margarine

level teaspoon salt

15 g flour

Total weight of ingredients 780 g.

Only ingredient with G + D (measured) is potatoes (400 g).

Content of potato in fish pie (%)

\[
\frac{400}{780} = 0·5128 \%
\]

G + D content of potato (old and new) = (0·74 + 3·75)/2 = 2·6 μg per 100 g.
G + D content for total fish pie = \((2.6 \times 51.28)/100\) µg per 100g
= 1.3 µg per 100g.

Assume 10.1% weight loss on cooking and G + D content rises by similar amount, then the G + D content of fish pie = 1.3 + \((10.1 \times 1.3)\) = \((1.3 + 0.13)\) µg per 100 g = 1.5 µg per 100 g.

The same mathematical formula was used to calculate the G + D content of other fish products such as fishcakes (see Table 5).

**Isoflavone (genistein plus daidzein) values for retail products**

Recipe for carrot and nut salad with French dressing; food code 15 288 (Tesco)

- 42 g carrot
- 35 g French dressing
- 15 g groundnuts
- 8 g sultanas
- Total weight 100 g.

Groundnut content (%) = 15; G + D content of groundnuts is 20.9 µg per 100 g.
Sultanas content (%) = 8; G + D content of sultanas is 204.1 µg per 100 g.

So, in 100 g salad G + D content = \(((15 \times 20.9)/100 + (8 \times 204.1)/100)\) µg G + D
= \((3.1 + 16.3)\) µg G + D

Phyto-oestrogen content of carrot and nut salad = 19.4 µg per 100 g.

**Appendix 2. Genistein plus daidzein content of Streaky Strips®**

Streaky Strips®, food code MR11, contain 43% textured vegetable protein (TVP; wheat gluten and soya protein concentrate), soya and maize protein, and soya protein isolate (SPI; assume 14%)

We can make the following assumptions:
- 50% TVP is wheat protein and 50% is soya protein;
- Genistein plus daidzein (G + D) content is 50% G + D content of soya flour (Riaz, ‘Soya and Health 2000’).

100 g Streaky Strips® contains 23 g TVP and 14 g SPI.

G + D content of TVP = 68,600 µg per 100 g.
G + D content of SPI = 105,000 µg per 100 g.

So, total G + D content = \(((23 \times 68,600) + 14 \times 105,000)/100\)

Phyto-oestrogen content of Streaky Strips® = \((15,778 + 14,700)\) µg per 100 g = 30,478 µg per 100 g.