Quality Control of Nutrient Data for a Longterm, Multi-Centre Dietary Intervention Trial

Cary Greenberg RD
Nutrition Coordinator,
Diet and Breast Cancer Prevention Study
Ontario Cancer Institute
Division of Epidemiology and Statistics

May 2008
Diet and Breast Cancer Prevention Study

- Canadian
- multi-centre
- in operation from 1988 – 2005
- randomized, intervention trial
Study Goal

to determine if a low fat, high carbohydrate diet will reduce the incidence of breast cancer in healthy women with extensive mammographic density.
Diet and Breast Cancer Prevention Study

Design

Eligible Subjects Identified

Prerandomization Assessment

Intervention (n=2,341)

Control (n=2,349)

Annual Visits

• demo/anthro data
• 3 day food records
• non fasting serum

Follow up until Dec 2005 (7-17 years per subject)
## Selected Baseline Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Intervention</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Subjects</td>
<td>2341</td>
<td>2349</td>
</tr>
<tr>
<td>Age (years)</td>
<td>47.1</td>
<td>47.2</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>62.2</td>
<td>62.3</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.3</td>
<td>23.3</td>
</tr>
<tr>
<td>Parity (% parous)</td>
<td>73.8</td>
<td>74.5</td>
</tr>
<tr>
<td>First degree relative with BC(%)</td>
<td>17.7</td>
<td>18.3</td>
</tr>
<tr>
<td>Premenopausal (%)</td>
<td>73.2</td>
<td>73.9</td>
</tr>
</tbody>
</table>
82,000 Food Record Days Analyzed

- In 6 sites
- Over 17 years
- By 46 dietitians or nutritionists
“Quality is never an accident, it is the result of intelligent effort.”
Quality Control Program

*for accuracy and consistency*

- Food Knowledge Test
- Extensive Training
- “Missing Food Book”
- Review Out of Range food records
- Double entry
Food Knowledge Test

1. Measurements
   tsp in a tbsp = tbsp in a cup= oz in a lb= FO in a cup=

2. Commercial and Ethnic Foods
   What is Orangina, Jamaican Patty, Challah, Marmite, Boursin, Mesclun Mix

3. Math Skills
   Calculate the volume of a wedge using this formula -
   \[ \frac{1}{2} \times \text{width of rounded edge} \times \text{radius} \times \text{height} \]
   rounded edge = 3” radius = 4” height = 2.5”

4. Nutrition Composition
   Approximately how many calories in 4 oz of cooked, lean beef?
NDS Training and Certification

- Close supervision for about 3 months.
- First 100 food records checked by Nutrient Data Manager.
- Results reviewed with staff and feedback given for all errors.
“Missing Food Book”

supplementary electronic database with information and guidelines for nutrient data entry

1. Product Information ~ 12,000 new foods
2. Data Entry Rules
3. Guides for Entry – variance guide, sushi, ribs, calculating volumes of soup
## Eating Habit Survey

### Missing Food Form

#### Product Name
- [ ] Grape
- [ ] Apple
- [ ] Orange
- [ ] Pear
- [ ] other

#### Information Source(s)
- [ ] Picture
- [ ] Telephone
- [ ] other

#### Package (enter all that apply)
- [ ] tranche
- [ ] pound
- [ ] each
- [ ] other

#### Preparation
- [ ] package
- [ ] other

#### Main Ingredients
- [ ]
- [ ]

#### Nutrients

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Per 100g</th>
<th>Per 200g</th>
<th>Per 300g</th>
<th>Per 400g</th>
<th>Per 500g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ca</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Notes

*Please attach any other notes here.*
Criteria for “Out of Range” Records

Energy $\leq$ 800 calories
$\geq$ 2800 calories

Fat $\geq$ 110 grams

Impact of “Out of Range” Records

- 10 – 12 % food records are out of range
- 35% records changed after review
- 65% false positive
Double Entry Project

“Quality is not an act, it is a habit.”

Aristotle
Double Entry Project

Goal:
To determine the reliability of data entry between dietitians and different versions of NDS.

Objective:
To compare the nutrient data obtained from records collected before 1996 and entered in early and later versions (2.91) of NDS.
Method

- Randomly select 50 sets from each study group (total 100 3-day record sets)
- Used Y2 records previously entered for hormone sub-study
- Distributed blinded records to all available dietitians
## Results

<table>
<thead>
<tr>
<th></th>
<th>Initial Entry Mean</th>
<th>Repeat Entry Mean</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy (kcal)</td>
<td>1618</td>
<td>1617</td>
<td>0.96</td>
</tr>
<tr>
<td>Total Fat (g)</td>
<td>49.0</td>
<td>48.8</td>
<td>0.96</td>
</tr>
<tr>
<td>Protein (g)</td>
<td>66.5</td>
<td>65.0</td>
<td>0.93</td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>227.7</td>
<td>230.4</td>
<td>0.98</td>
</tr>
</tbody>
</table>
## Nutrient Analysis Results

### Intake of selected nutrients by group and time

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean (SD)</th>
<th>Baseline</th>
<th>Year 1 - 2</th>
<th>Year 4 - 6</th>
<th>Year 8 - 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy (Kcal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1690 (361)</td>
<td>1642 (349)</td>
<td>1612 (356)</td>
<td>1533 (341)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1739 (377)</td>
<td>1753 (386)</td>
<td>1757 (390)</td>
<td>1738 (403)</td>
<td></td>
</tr>
<tr>
<td>Total fat (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>28.6 (6.4)</td>
<td>18.8 (5.7)</td>
<td>20.5 (6.4)</td>
<td>22.0 (6.2)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>29.7 (6.7)</td>
<td>29.7 (6.9)</td>
<td>29.8 (6.4)</td>
<td>31.2 (7.3)</td>
<td></td>
</tr>
<tr>
<td>Total CHO (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>54.7 (7.5)</td>
<td>64.7 (7.1)</td>
<td>62.4 (8.1)</td>
<td>59.9 (8.0)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>53.2 (7.6)</td>
<td>53. (8.0)</td>
<td>52.8 (8.1)</td>
<td>50.8 (9.0)</td>
<td></td>
</tr>
<tr>
<td>Total Protein (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>16.3 (3.1)</td>
<td>17.0 (2.7)</td>
<td>17.3 (3.0)</td>
<td>18.5 (3.9)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>16.0 (2.8)</td>
<td>16.2 (3.2)</td>
<td>16.6 (3.1)</td>
<td>17.2 (3.3)</td>
<td></td>
</tr>
</tbody>
</table>

I = intervention (n=358)
C = comparison (n=365)
Body weight by group and time of follow-up

Baseline | Years 1-2 | Years 4-6 | Years 8-10

Comparison (n=365)

Intervention (n=358)
Blood cholesterol levels by group and time of follow-up (adjusted for age and weight)

Total Cholesterol (mmol/l)  HDL (mmol/l)
“I am easily satisfied with the best.”

Winston Churchill