Alkylresorcinols in Adipose Tissue as Long-term Biomarkers of Whole Grain Wheat and Rye Intake

Huaxing Wu
Faculty of Natural Resources and Agricultural Sciences
Department of Molecular Sciences
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2017
Cover: Whole grain rye and wheat, from blurred to clear
(Photo: Wenbo Dong)
Alkylresorcinols (AR) are phenolic lipids present in the outer parts of rye and wheat grain kernels. Total AR concentration and the ratio of two AR homologs, C17:0/C21:0, in plasma have been evaluated and used as biomarkers of whole grain rye and wheat intake. The AR C17:0/C21:0 ratio is 1.0 in rye and 0.1 in common wheat and reflects the relative intake of whole grain rye. However, due to the short half-life (4-5 h) of AR in plasma, more long-term biomarkers are needed.

This thesis investigated whether AR in adipose tissue could be used as long-term biomarkers of whole grain wheat and rye intake, and whether these biomarkers, in parallel with whole grain intakes determined by dietary assessment, were associated with blood lipid profiles and breast cancer incidence in two Nordic populations.

A sensitive, high-throughput GC-MS method for quantification of AR concentrations in adipose tissue biopsies was developed, evaluated, and applied. The AR concentrations in adipose tissue were evaluated as long-term biomarkers by correlating them to whole grain intake, measured precisely by repeated food records during a 12-week intervention, and to long-term whole grain wheat and rye intake estimated with food frequency questionnaires (FFQs) in observational studies. Intake was found to be modestly to well correlated with total AR concentrations in adipose tissue, with sex being an important determinant. The performance of AR in adipose tissue as biomarkers of whole grain wheat and rye appeared comparable to that of AR in plasma, and therefore they were further tested in relation to blood lipid profiles and incidence of breast cancer in two observational studies.

Overall, total whole grain wheat and rye, reflected by biomarkers or dietary assessment, were not associated with blood lipids or breast cancer incidence in the populations studied. However, AR C17:0/C21:0 ratio in adipose tissue was inversely associated with low-density lipoprotein cholesterol (LDL-C) in blood and positively associated with the risk of developing ER+ breast cancer among women. These findings suggest that whole grain wheat and rye may have different effects on the endpoints studied but need to be confirmed in other population. The results illustrate the importance of using dietary biomarkers in parallel with conventional dietary assessment.

**Keywords:** Adipose tissue, alkylresorcinols, biomarkers, breast cancer, cholesterol, plasma, rye, wheat, whole grain

**Author’s address:** Huaxing Wu, Department of Molecular Sciences, SLU P.O. Box 7015, 750 07 Uppsala, Sweden
Dedication

To my parents

五谷为养
Grains and beans, backbone of nutrition
-黄帝内经
-Yellow Emperor’s Classic of Internal Medicine
**Contents**

List of publications .......................................................... 7

Abbreviations ...................................................................... 10

1  **Background** .................................................................. 13
   1.1 Whole grains and health .............................................. 14
      1.1.1 Whole grains and cardiovascular disease ............... 16
      1.1.2 Whole grains and breast cancer ......................... 18
      1.1.3 The potential role of individual whole grains on health 20
   1.2 Estimation of whole grain intake .............................. 20
   1.3 Alkylresorcinols in cereals ........................................ 21
   1.4 Absorption, distribution, and elimination of alkylresorcinols 23
      1.4.1 Absorption .......................................................... 23
      1.4.2 Distribution ......................................................... 23
      1.4.3 Elimination .......................................................... 23
   1.5 Alkylresorcinols and their metabolites as dietary biomarkers 25
      1.5.1 Dietary biomarkers .............................................. 25
      1.5.2 Alkylresorcinols in plasma and their metabolites in plasma or urine as intake biomarkers 27
      1.5.3 Alkylresorcinols in adipose tissue as long-term biomarkers? 29

2  **Aims of thesis** ............................................................... 30

3  **Materials and methods** ................................................ 31
   3.1 Reference compounds .............................................. 31
   3.2 Instrumentation ......................................................... 31
   3.3 Alkylresorcinols analysis in adipose tissue and plasma .... 31
      3.3.1 Sample extraction .............................................. 33
      3.3.2 Solid phase extraction clean-up ............................ 33
      3.3.3 Derivatization ..................................................... 34
      3.3.4 Analysis of alkylresorcinols on GC-MS ................. 34
      3.3.5 Calibration .......................................................... 35
      3.3.6 Evaluation of the method .................................... 35
   3.4 Samples and study designs ........................................ 36
3.4.1 Populations, subjects and samples 36
3.4.2 Dietary assessment of whole grain intake 40
3.5 Statistical analysis 42

4 Results and discussion 45
4.1 Determination of alkylresorcinols in adipose tissue 45
4.2 Alkylresorcinols in adipose tissue in response to a 12-week controlled whole grain intervention 51
4.3 Alkylresorcinols in adipose tissue as biomarkers under free-living conditions 53
4.4 Alkylresorcinols in adipose tissue vs. alkylresorcinols in plasma as biomarkers 57
4.5 Alkylresorcinols in adipose tissue in relation to blood lipids 58
4.6 Alkylresorcinols in adipose tissue in relation to breast cancer incidence 60

5 General discussion 63

6 Main findings 66

7 Future research 67

References 68

Acknowledgements 83
List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:


IV Wu, H.*, Omar, NA., Håkansson, N., Wolk, A., Michaëlsson, K. & Landberg, R. Long-term whole-grain rye but not wheat intake reflected by alkylresorcinols in adipose tissue, is associated with lower LDL-cholesterol in Swedish women. *(Manuscript)*

Papers I-II are reproduced with the permission of the publishers.

*Corresponding author

Associated papers not included in this doctoral thesis:


The contribution of Huaxing Wu to Papers I-V was as follows:

I: Participated in planning the experiments, performed the laboratory and statistical analyses of alkylresorcinol concentrations in plasma and adipose tissue, and had shared main responsibility for writing and revising the manuscript with Roksana Wierzbicka (shared first authorship).

II: Performed the statistical analyses and had main responsibility for writing and revising the manuscript.

III: Participated in planning experiments, performed the laboratory- and statistical analyses, and had main responsibility for writing and revising the manuscript.

IV: Planned the study in collaboration with supervisors, performed the laboratory and statistical analyses, and had main responsibility for writing and revising the manuscript.

V: Planned the study in collaboration with supervisors, performed the laboratory and statistical analyses, and had main responsibility for writing and revising the manuscript.
Abbreviations

AACC  American Association of Cereal Chemists
ANOVA  Analysis of variance
AR  Alkylresorcinol(s)
AX  Arabinoxylan
CI  Confidence interval
COSM  Cohort of Swedish Men
CVD  Cardiovascular disease
EI  Electron impact ionization
ER  Estrogen receptor
ER-  Estrogen receptor negative
ER+  Estrogen receptor positive
FFQ  Food frequency questionnaire
GC  Gas chromatography
GC  Gas chromatography- mass spectrometry
GLM  Generalized linear model
HDL-C  High-density lipoprotein-cholesterol
HR  Hazard ratios
HRT  Hormone replacement therapy
ICC  Intra-class correlation
LDL-C  Low-density lipoprotein-cholesterol
LOD  Limit of detection
LOQ  Limit of quantification
MUFA  Monounsaturated fatty acids
NCI  Negative chemical ionization
PUFA  Polyunsaturated fatty acids
QSM  Quick silylation mixture
RDG  Refined diet group
SFA  Saturated fatty acids
SIM  Selected ion monitoring
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMC</td>
<td>Swedish Mammography Cohort</td>
</tr>
<tr>
<td>SPE</td>
<td>Solid phase extraction</td>
</tr>
<tr>
<td>TCHOL</td>
<td>Total cholesterol level</td>
</tr>
<tr>
<td>TFAA</td>
<td>Trifluoroacetic anhydride</td>
</tr>
<tr>
<td>TG</td>
<td>Triglyceride</td>
</tr>
<tr>
<td>TIC</td>
<td>Total ion chromatogram</td>
</tr>
<tr>
<td>TWG</td>
<td>Total whole grain intake from FFQ</td>
</tr>
<tr>
<td>WGDG</td>
<td>Whole grain diet group</td>
</tr>
<tr>
<td>WGR</td>
<td>Whole grain rye intake from FFQ</td>
</tr>
<tr>
<td>WGR%</td>
<td>Relative whole grain rye intake from FFQ</td>
</tr>
<tr>
<td>WGR&amp;W</td>
<td>Whole grain rye and wheat intake from FFQ</td>
</tr>
<tr>
<td>WGW</td>
<td>Whole grain wheat intake from FFQ</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
1 Background

Non-communicable diseases are major threats to sustainable development globally and are estimated to have account for around 60% of all deaths in 2002 (WHO, 2005), 63% in 2008 (WHO, 2013), and 70% in 2015 (WHO, 2017), while the number is projected to increase to 73% by 2020 (WHO, 2005). Cardiovascular disease (CVD) (17.7 million deaths) and cancer (9.0 million deaths) are the two leading causes of death world-wide (WHO, 2017).

CVD is strongly affected by lifestyle (Lichtenstein et al., 2006). A number of behavioral risk factors, including alcohol and tobacco use, physical inactivity, an unhealthy diet with excess salt, fat, and energy, and metabolic risk factors, such as elevated blood pressure, blood sugar, blood lipids, and body weight, play major roles in development of CVD (Mendis et al., 2011). Despite high mortality rates, CVD mortality has been significantly reduced in high-income countries (Western Europe, North America, Australia, New Zealand, Japan, and the Republic of Korea), due to promotion of a healthy lifestyle and good healthcare. An increase in mortality is mainly seen in middle- and low-income countries, where it is related to an unhealthy lifestyle and rising life expectancy (WHO, 2014).

Cancer is also highly affected by lifestyle factors, including diet (Willett, 2002, Willett et al., 2006). Breast cancer is the most common cancer among women (25% of all cancers in female) with 1.7 million new cases and 0.5 million deaths in 2012 (Stewart and Wild, 2014). Breast cancer risk factors include age (higher at younger ages), age at menarche and menopause (increased risk for earlier menarche and/or later menopause ), age at first pregnancy (risk of breast cancer for women have their first child at age: before 20 < after 30 < after 35), family history (including BRCA1, BRCA2, and other genes), previous benign breast disease (severe atypical epithelial hyperplasia), exposed to radiation (mainly during the World War II), use of oral contraceptives, hormone replacement therapy (HRT), use of alcohol (McPherson et al., 2000), and infection with human mammary tumor virus.
Based on expression of estrogen receptor, breast cancer can be divided into estrogen-receptor-positive (ER+) and estrogen-receptor-negative (ER-) subsets, where the former responds better to endocrine therapy and lifestyle change (Rugo et al., 2016).

1.1 Whole grains and health

Much attention has been devoted to the role of nutrients in prevention of non-communicable disease and reducing mortality in the past (Garland and Garland, 1980, Trowell, 1972). More recently, attention has turned to the role of foods and food patterns to a greater extent (Hu, 2002).

Cereals are important staple foods and are the main source of energy intake in the US and Europe (O’Neil et al., 2012, Ocke et al., 2009). During the past 100 years, cereals have generally been consumed as refined grains, which mainly consist of starchy endosperm. Whole grain, on the other hand, contains all parts of the grain, i.e., the starchy endosperm, bran, and germ, and is a rich source of nutrients (Table 1), including vitamins (Melanson et al., 2006), dietary fibers (Turner and Lupton, 2011), unsaturated fatty acids (Slavin, 2003), and phytochemicals (Liu, 2007), and minerals (Melanson et al., 2006).

Accumulating data, mainly from observational studies, suggests beneficial health effects of whole grain foods, and increased intake is therefore recommended by national dietary guidelines in many countries (Slavin et al., 2016). For example, in the latest Nordic Nutrition Recommendations 2012, refined cereals are recommended to be replaced with whole grain cereals to reduce disease risk and sustain health (Becker et al., 2013).

A common definition of whole grain that has been used in many countries, including in the Nordic Nutrition Recommendations, is that provided by the American Association of Cereal Chemists (AACC) from 1999:

“Whole grains shall consist of the intact, ground, cracked or flaked caryopsis, whose principal anatomical components—the starchy endosperm, germ, and bran—are present in the same relative proportions as they exist in the intact caryopsis.”

The AACC definition include caryopses from a broad range of grains classified into three groups: (1) Cereals: wheat, triticale, rye, corn, barley, oats, millet, sorghum, and rice; (2) pseudocereals, including amaranth, buckwheat and quinoa; and (3) wild rice.
### Table 1. Nutrient composition (per 100 g) of refined wheat and common whole grains cereals

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Unit</th>
<th>Refined wheat</th>
<th>Wheat</th>
<th>Rye</th>
<th>Oat</th>
<th>Buckwheat</th>
<th>Corn</th>
<th>Barley</th>
<th>Brown rice</th>
<th>Millet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>kcal</td>
<td>361</td>
<td>340</td>
<td>325</td>
<td>404</td>
<td>335</td>
<td>361</td>
<td>345</td>
<td>363</td>
<td>382</td>
</tr>
<tr>
<td>Protein</td>
<td>g</td>
<td>12.0</td>
<td>13.2</td>
<td>15.9</td>
<td>14.7</td>
<td>12.6</td>
<td>6.9</td>
<td>10.5</td>
<td>7.2</td>
<td>10.8</td>
</tr>
<tr>
<td>Total lipids</td>
<td>g</td>
<td>1.7</td>
<td>2.5</td>
<td>2.2</td>
<td>9.1</td>
<td>3.1</td>
<td>3.9</td>
<td>1.6</td>
<td>2.8</td>
<td>4.3</td>
</tr>
<tr>
<td>SFA(^2)</td>
<td>g</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
<td>1.6</td>
<td>0.7</td>
<td>0.5</td>
<td>0.3</td>
<td>0.6</td>
<td>0.5</td>
</tr>
<tr>
<td>MUFA(^2)</td>
<td>g</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>2.9</td>
<td>0.9</td>
<td>1.0</td>
<td>0.2</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td>PUFA(^2)</td>
<td>g</td>
<td>0.7</td>
<td>1.2</td>
<td>1.0</td>
<td>3.3</td>
<td>0.9</td>
<td>1.8</td>
<td>0.8</td>
<td>1.0</td>
<td>2.6</td>
</tr>
<tr>
<td>Carbohydrate</td>
<td>g</td>
<td>72.5</td>
<td>72.0</td>
<td>68.6</td>
<td>65.7</td>
<td>70.6</td>
<td>76.9</td>
<td>74.5</td>
<td>76.5</td>
<td>75.1</td>
</tr>
<tr>
<td>Dietary fiber</td>
<td>g</td>
<td>2.4</td>
<td>10.7</td>
<td>23.8</td>
<td>6.5</td>
<td>10.0</td>
<td>7.3</td>
<td>10.1</td>
<td>4.6</td>
<td>3.5</td>
</tr>
<tr>
<td>Sugars</td>
<td>g</td>
<td>0.3</td>
<td>0.4</td>
<td>2.3</td>
<td>0.8</td>
<td>2.6</td>
<td>0.6</td>
<td>0.8</td>
<td>0.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Calcium</td>
<td>mg</td>
<td>15</td>
<td>34</td>
<td>37</td>
<td>55</td>
<td>41</td>
<td>7</td>
<td>32</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Iron</td>
<td>mg</td>
<td>0.9</td>
<td>3.6</td>
<td>5.0</td>
<td>4.0</td>
<td>4.1</td>
<td>2.4</td>
<td>2.7</td>
<td>2.0</td>
<td>3.9</td>
</tr>
<tr>
<td>Magnesium</td>
<td>mg</td>
<td>25</td>
<td>137</td>
<td>160</td>
<td>144</td>
<td>251</td>
<td>93</td>
<td>96</td>
<td>112</td>
<td>119</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>mg</td>
<td>97</td>
<td>357</td>
<td>499</td>
<td>452</td>
<td>337</td>
<td>272</td>
<td>296</td>
<td>337</td>
<td>285</td>
</tr>
<tr>
<td>Potassium</td>
<td>mg</td>
<td>100</td>
<td>363</td>
<td>717</td>
<td>371</td>
<td>577</td>
<td>315</td>
<td>309</td>
<td>289</td>
<td>224</td>
</tr>
<tr>
<td>Sodium</td>
<td>mg</td>
<td>2</td>
<td>2</td>
<td>19</td>
<td>11</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Zinc</td>
<td>mg</td>
<td>0.9</td>
<td>2.6</td>
<td>5.0</td>
<td>3.2</td>
<td>3.1</td>
<td>1.7</td>
<td>2.0</td>
<td>2.5</td>
<td>2.6</td>
</tr>
<tr>
<td>Thiamin</td>
<td>mg</td>
<td>0.08</td>
<td>0.50</td>
<td>0.32</td>
<td>0.69</td>
<td>0.42</td>
<td>0.25</td>
<td>0.37</td>
<td>0.44</td>
<td>0.41</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>mg</td>
<td>0.06</td>
<td>0.17</td>
<td>0.25</td>
<td>0.13</td>
<td>0.19</td>
<td>0.08</td>
<td>0.11</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Niacin</td>
<td>mg</td>
<td>1.0</td>
<td>5.0</td>
<td>4.3</td>
<td>1.5</td>
<td>6.2</td>
<td>1.9</td>
<td>6.3</td>
<td>6.3</td>
<td>6.0</td>
</tr>
<tr>
<td>Vitamin B-6</td>
<td>mg</td>
<td>0.04</td>
<td>0.17</td>
<td>0.25</td>
<td>0.13</td>
<td>0.58</td>
<td>0.37</td>
<td>0.40</td>
<td>0.74</td>
<td>0.37</td>
</tr>
<tr>
<td>Folate</td>
<td>µg</td>
<td>33</td>
<td>44</td>
<td>33</td>
<td>32</td>
<td>54</td>
<td>25</td>
<td>8</td>
<td>16</td>
<td>42</td>
</tr>
<tr>
<td>Vitamin E</td>
<td>mg</td>
<td>0.4</td>
<td>0.7</td>
<td>2.7</td>
<td>0.7</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>Vitamin K</td>
<td>µg</td>
<td>0.3</td>
<td>1.9</td>
<td>5.9</td>
<td>3.2</td>
<td>7</td>
<td>0.3</td>
<td>2.2</td>
<td>0.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

\(^1\) Data from the National nutrient database for standard reference Release 28 (USDA, 2017).  
\(^2\) SFA, saturated fatty acids; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid
However, there is currently no widely accepted definition of a whole grain product. The HEALTHGRAIN Forum recently proposed a definition of a whole grain product as: “A whole-grain food is one for which the product is made with $\geq 30\%$ whole-grain ingredients on a dry-weight basis and more whole-grain ingredients than refined-grain ingredients” (Ross et al., 2017b). From a scientific point of view, the amount of whole grain consumed per day may be of greatest relevance in relation to health outcomes (Aune et al., 2016). It is therefore crucial that whole grain consumption can be estimated accurately.

The evidence of beneficial health effects of whole grains used to justify the dietary guidelines is primarily based on results from observational studies. Observational studies have consistently shown that high whole grain intake is inversely associated with risk of CVD (Aune et al., 2016, Jacobs and Gallaher, 2004, Mellen et al., 2008, Wu et al., 2015), type 2 diabetes (Cho et al., 2013, Meyer et al., 2000, Salmerón et al., 1997, Ye et al., 2012), and colorectal cancer (Aune et al., 2016, Chatenoud et al., 1998, Jacobs et al., 1998). There is yet a number of outcomes where a protective role of whole grains has been suggested and investigated based on plausible hypotheses, but where results remain inconclusive. Breast cancer is one such outcome.

### 1.1.1 Whole grains and cardiovascular disease

A recent meta-analysis suggested that three daily servings of whole grains was associated with a 19, 12 and 22% lower risk of developing coronary heart disease, stroke and total CVD, respectively (Aune et al., 2016). Despite large differences in the types of grains consumed in different parts of the world (Table 2), studies included in the meta-analysis (Aune et al., 2016) consistently show protective associations for a wide range of types and amount of grains consumed (Table 3).

**Table 2.** The most commonly consumed whole grains from different parts of the world reported in some studies.

<table>
<thead>
<tr>
<th>Country</th>
<th>Whole grain</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>Wheat (14-19%), rye (53-56%), and oats (15-16%)</td>
<td>(Kyrø et al., 2012)</td>
</tr>
<tr>
<td>Norway</td>
<td>Wheat (72%), rye (20%), and oats (6%)</td>
<td>(Kyrø et al., 2012)</td>
</tr>
<tr>
<td>Denmark</td>
<td>Wheat (8-10%), rye (71-73%), and oats (15-19%)</td>
<td>(Kyrø et al., 2012)</td>
</tr>
<tr>
<td>US</td>
<td>Wheat, corn, oats, barley and rice</td>
<td>(Jonnalagadda et al., 2011)</td>
</tr>
<tr>
<td>France</td>
<td>Wheat, oats and buckwheat</td>
<td>(Bellisle et al., 2014)</td>
</tr>
<tr>
<td>China</td>
<td>Corn and sorghum</td>
<td>(Wang et al., 2016)</td>
</tr>
<tr>
<td>Region</td>
<td>n</td>
<td>Quantity of intake</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>US</td>
<td>75 521 women</td>
<td>1.77 serving/d vs. less than 0.26 serving/d of whole grain</td>
</tr>
<tr>
<td>US</td>
<td>75 521 women</td>
<td>1.77 serving/d vs. less than 0.26 serving/d</td>
</tr>
<tr>
<td>US</td>
<td>75 521 women and 43 744 men</td>
<td>Highest quintile with median 33.0 g/d for women and 47.8 g/d for men</td>
</tr>
<tr>
<td>US</td>
<td>566 399 men and women</td>
<td>Higher intake quintile with median 34.0 g/d vs. lowest quintile with median 3.7 g/d</td>
</tr>
<tr>
<td>US</td>
<td>535 subjects</td>
<td>Higher intake quintile with median 2.9 serving/d vs. lowest quintile with median 0.31 serving/d</td>
</tr>
<tr>
<td>US</td>
<td>11 940 subjects</td>
<td>Higher intake quintile with median 3 serving/d vs. lowest quintile with median 0.1 serving/d</td>
</tr>
<tr>
<td>US</td>
<td>34 827 women</td>
<td>Quintiles of whole grain divided by DASH Score: higher intake quintile with mean 2.33 serving/d vs. lowest quintile with mean 0.78 serving/d</td>
</tr>
<tr>
<td>Norway, Sweden and Denmark</td>
<td>120 010 men and women</td>
<td>Per doubling in intake</td>
</tr>
<tr>
<td>Sweden</td>
<td>32 561 women</td>
<td>Per 90g/day increase</td>
</tr>
<tr>
<td>Finland</td>
<td>3 932 men and women</td>
<td>Highest quintile 195-963 g/d for women and 280-1321 g/d for men vs. lowest quintile 0-89 g/d for women and 0-139 g/d for men</td>
</tr>
<tr>
<td>Spain</td>
<td>7 447 men and women</td>
<td>Higher intake quintile with median 84 g/d vs. lowest quintile with mean 0 g/d</td>
</tr>
<tr>
<td>China</td>
<td>1 104 men and 1 341 women</td>
<td>Per once/d</td>
</tr>
</tbody>
</table>
Potential mediators of beneficial effects of whole grains on cardiovascular disease

The observed protective role of whole grains on CVD might be mediated through effects on body weight, blood pressure, glycemic control, blood lipids and inflammation which are all recognized as cardiometabolic risk factors (Behall et al., 2006, Dudina et al., 2011, Esmaillzadeh et al., 2004, Jensen et al., 2006, McKeown et al., 2002, Montonen et al., 2013, Qi et al., 2006, Tighe et al., 2010).

One of the most established mechanism of protective effects of whole grains on CVD is through total cholesterol (TCHOL) and LDL-C reduction caused by viscous dietary fiber and phytosterols (Hunninghake et al., 1994, Ripsin et al., 1992). The cholesterol lowering effect has been shown for beta-glucans from oats and barley and there are approved health claims in Europe (Agostoni et al., 2011). Whole grain foods are good source of phytosterols (Ostlund and Lin, 2006) and intakes of 2.3 ± 0.5 g/d have been shown to lower TCHOL and LDL-C by 7-11% and 10-15%, respectively (Moruisi et al., 2006).

Whole grain foods with well-preserved structure and a high content of viscous fiber could also delay the absorption of macronutrients, leading to better glycemic control (Cara et al., 1992, Hallfrisch et al., 1995). Moreover, whole grain intake could reduce total energy intake, as well as increase short-chain fatty acid production from gut microbiota (Clark and Slavin, 2013, Cooper et al., 2015). However, the joint or separate role of whole grain wheat and rye for most of the mechanisms above remains yet unclear due to few studies (Hollaender et al., 2015, Nilsson et al., 2008, Rosén et al., 2011, Slavin et al., 1999, Ye et al., 2012).

1.1.2 Whole grains and breast cancer

Although whole grain intake has been hypothesized to have a protective role for development of breast cancer (Jacobs et al., 1995), the overall association between intake of whole grain or cereal fiber and risk of breast cancer is inconsistent and may be affected by the use of hormone replacement therapy and menopausal status (Table 4).

Suggested mechanisms for impact of whole grain intake on breast cancer

Whole grain-derived dietary fiber and lignans may play roles in preventing the development of breast cancer (Adlercreutz, 2010, Jacobs et al., 1995, Slavin, 2000). Dietary fiber can interfere with estrogen metabolism and interrupt reabsorption of estrogens and their precursors in the intestine (Arts et al., 1991a, Arts et al., 1992, Braaten et al., 1994), while plant lignans can be metabolized by the mammalian microflora to enterolactone and enterodiol, which have
<table>
<thead>
<tr>
<th>Region</th>
<th>n</th>
<th>Quantity of Intake</th>
<th>Outcomes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>25 278</td>
<td>Whole grain product intake: highest quintile (more than 163 g/d) vs lowest quintile (less than 72 g/d)</td>
<td>Not significant</td>
<td>(Egeberg et al., 2009)</td>
</tr>
<tr>
<td>Greece</td>
<td>500</td>
<td>Whole grain intake: more than 7 times/week vs Never/rarely</td>
<td>51% lower breast incidence</td>
<td>(Mourouti et al., 2016)</td>
</tr>
<tr>
<td>US</td>
<td>90 534</td>
<td>Whole grain intake: highest quintile (44.8±16.8 g/d) vs lowest quintile (2.8±2.0 g/d)</td>
<td>26% lower premenopausal breast cancer incidence</td>
<td>(Farvid et al., 2016)</td>
</tr>
<tr>
<td>US</td>
<td>29 119</td>
<td>Whole grain intake: highest quintile (19-108.5 g/d) vs lowest quintile (0-3.5 g/d)</td>
<td>Not significant</td>
<td>(Nicodemus et al., 2001)</td>
</tr>
<tr>
<td>Denmark, France, Germany, Greece, Italy, Norway, Spain, Sweden, the Netherlands, and UK</td>
<td>334 849</td>
<td>Cereal fiber intake: highest quintile (more than 10.9 g/d) vs lowest quintile (less than 4.7 g/d)</td>
<td>Not significant</td>
<td>(Ferrari et al., 2013)</td>
</tr>
<tr>
<td>UK</td>
<td>35 792</td>
<td>Cereal fiber intake: highest quintile (more than 13 g/d) vs lowest quintile (less than 4 g/d)</td>
<td>Not significant</td>
<td>(Cade et al., 2007)</td>
</tr>
<tr>
<td>France</td>
<td>4 684</td>
<td>Cereal fiber intake: highest quintile (more than 7.7 g/d) vs lowest quintile (less than 4.4 g/d)</td>
<td>Not significant</td>
<td>(Deschasaux et al., 2013)</td>
</tr>
<tr>
<td>Canada</td>
<td>89 835</td>
<td>Cereal fiber intake: highest quintile (more than 5.6 g/d) vs lowest quintile (less than 2.6 g/d)</td>
<td>Not significant</td>
<td>(Terry et al., 2002)</td>
</tr>
<tr>
<td>US</td>
<td>88 678</td>
<td>Cereal fiber intake: highest quintile (median 8.4 g/d) vs lowest quintile (median 2.4 g/d)</td>
<td>Not significant</td>
<td>(Holmes et al., 2004)</td>
</tr>
<tr>
<td>US</td>
<td>116 671</td>
<td>Cereal fiber intake: highest quintile (median 8.8 g/d) vs lowest quintile (median 4.2 g/d)</td>
<td>Not significant</td>
<td>(Cho et al., 2003)</td>
</tr>
<tr>
<td>US</td>
<td>185 598</td>
<td>Cereal fiber intake: highest quintile (median 8.9 g/d) vs lowest quintile (median 2.5 g/d)</td>
<td>Not significant</td>
<td>(Park et al., 2009)</td>
</tr>
<tr>
<td>Sweden</td>
<td>51 823</td>
<td>Cereal fiber intake: highest quintile (median 21.5 g/d) vs lowest quintile (median 10.4 g/d)</td>
<td>50% lower incidence among former use of postmenopausal hormone</td>
<td>(Suzuki et al., 2008)</td>
</tr>
</tbody>
</table>

1.1.3 The potential role of individual whole grains on health

Most of the epidemiological studies on whole grain and health, have typically investigated the role of total whole grain intake, whereas the specific role of different whole grains have been less investigated (Kyro et al., 2013, Kyro et al., 2014b). Grains differ considerably in their content and composition of nutrients and bioactive compounds (Table 1) and the type of grains consumed varies worldwide (van der Kamp et al., 2014) (Table 2). Among the different whole grains, wheat and rye are most commonly consumed in the Nordic countries (Engeset et al., 2015), but their separate role, particularly the role of rye, remains yet unclear due to few studies (Hollaender et al., 2015, Nilsson et al., 2008, Rosén et al., 2011, Slavin et al., 1999, Ye et al., 2012).

Compared with whole grain wheat, whole grain rye contains higher amounts of total and viscosity forming dietary fiber (Andersson et al., 2013, Hansen et al., 2003, Saastamoinen et al., 1989) and phytosterols (1098-1420 µg/g in rye vs. 670-959 µg/g in wheat). Whole grain rye may therefore play a more important role for cholesterol lowering effects than whole grain wheat. Moreover, whole grain rye, but not other cereals to the same extent, is a rich source of lignans that are converted to enterolactone and enterodiol that have estrogenic and antiestrogenic effects (Andersson et al., 2013, Nurmi et al., 2008). Based on this, rye has been hypothesized to reduce risk of breast cancer (Adlercreutz, 2010).

1.2 Estimation of whole grain intake

In observational studies, whole grain intake is typically estimated by self-reporting methods such as food frequency questionnaires (FFQ) and dietary recalls, whereas food records are typically used in dietary intervention studies (Thompson and Subar, 2008). These methods are prone to relatively large measurement errors, particularly FFQs, which are typically used to estimate long-term whole grain intake (Kipnis et al., 2002, Kristal et al., 2005). The measurement errors are both random and systematic, and result from misreporting, poor information from databases on whole grain products, variations in the whole grain content in different whole grain products, and intake variation over time (Willett, 2012). In questionnaires, typically only a few questions can be used to capture whole grain intake, and the questions may cover foods that are typically based on several grains. Moreover, it is difficult to estimate whole grain intake from different grains separately by the dietary
assessment methods, which might be one major reason why so few observational studies have been conducted on the separate role of different whole grains, while dietary intervention studies to date have generally used whole grains foods based on mixed grains (Brownlee et al., 2010, Kristensen et al., 2012, Tighe et al., 2010).

To overcome some of these obstacles and facilitate investigation of the role of different whole grain types in health, dietary biomarkers that provide objective measurement of specific whole grain intakes may be used (Landberg et al., 2014). Alkylresorcinols in plasma and their metabolites in plasma and urine have been suggested, evaluated, and applied as concentration biomarkers of whole grain wheat and rye intake (Ross, 2012).

1.3 Alkylresorcinols in cereals

Whole grain wheat, rye, barley (Kulawinek et al., 2008, Ross et al., 2003c), and quinoa (Ross et al., 2017a) contain 5n-alkylresorcinols, which are a group of amphiphilic phenolic lipids. In whole grains, alkylresorcinols (AR) content is highest in rye (720-761 µg/g) (Nyström et al., 2008, Ross et al., 2003c) and common wheat (410-761 µg/g) (Andersson et al., 2008a, Ross et al., 2003c), and significantly lower in barley (30-103 µg/g) (Andersson et al., 2008b, Ross et al., 2003c) and quinoa (32-93 µg/g) (Ross et al., 2017a). The main AR homologs present in wheat and rye consist of alkyl chains with odd number carbon atoms (C15:0, C17:0, C19:0, C21:0, C23:0 and C25:0) (Figure 1) (Kulawinek et al., 2008, Ross et al., 2003c) and are located in the outer part of the kernel (outer cuticula of the testa) (Landberg et al., 2008b). Recently, quinoa has been found to contain AR with even number carbon atoms chain or branched alkyl tails (Ross et al., 2017a). Unsaturated or keto/hydroxyl group substituted alkyl tails are also present in minor amounts in wheat and rye (Knödler et al., 2010, Kozubek and Tyman, 1995, Kozubek and Tyman, 1999, Seitz, 1992, Wieringa, 1967).

During milling, AR are mainly found in the bran fraction, leaving very low amounts in the starchy endosperm and germ (Kulawinek et al., 2008, Landberg et al., 2008b, Ross et al., 2003c). For this reason, bran and whole grain wheat and rye products are the main sources of AR in the Nordic population, while refined wheat only provides a minor contribution. The AR content in different whole grain kernels is influenced by the plant species, thousand-grain weight (usually a negative correlation), breeding, growth environment, and extraction ratio in milling (Andersson et al., 2008b, Ross et al., 2003c, Shewry et al., 2010). AR are stable and can resist common food processing method, as shown by good correlations between the whole grain content in different food product
ingredients and final products (Chen et al., 2004, Landberg et al., 2006a, Menzel et al., 2012, Ross et al., 2003c).

Figure 1. Most common whole grain wheat and rye derived alkylresorcinols with odd number carbon atoms alkyl chains from C17:0 to C25:0. Modified from (Landberg et al., 2014).

The relative homolog composition of AR varies between grains, but remarkably little within grain species (Andersson et al., 2008a, Ross et al., 2017a, Shewry et al., 2010). AR with both even and odd number carbon atoms alkyl chain are present in quinoa (Ross et al., 2017a), while the alkyl chains of AR found in rye and wheat only consist of odd number carbon atoms (Andersson et al., 2008a, Shewry et al., 2010). The AR C17:0/C21:0 ratio has been shown to differ between grains, with a ratio of 0.1 in whole grain wheat, 1.0 in whole grain rye, and 0.01 in durum wheat (Chen et al., 2004, Menzel et al., 2012, Ross et al., 2003c).
1.4 Absorption, distribution, and elimination of alkylresorcinols

1.4.1 Absorption

The absorption of AR was first studied in rats and about 50-65% of AR was found to be absorbed (Tłuskik et al., 1990). Ross and co-workers showed in an ileostomy operated pig model that AR are absorbed in the upper part of the gastrointestinal tract, with about 60-80% of the ingested dose absorbed (Ross et al., 2003b) and that about 60% was absorbed in 10 ileostomy operated subjects (Ross et al., 2003a).

In common with other lipophilic compounds, AR are not soluble in blood and therefore cannot be easily injected and therefore the absolute bioavailability has not been determined (Landberg et al., 2006b). However, higher relative bioavailability for longer AR homologs was suggested based on differences in dose-adjusted areas under the plasma concentration-time profiles for different homologs (Landberg et al., 2006b).

1.4.2 Distribution

After absorption in the small intestine, AR are transported to the systemic circulation either through the lymphatic pathway or portal vein (Linko-Parvinen et al., 2007, Marklund et al., 2014b). There is reported no differences in portal vein blood and arterial blood AR concentrations, suggesting that lymphatic absorption is the main route (Landberg et al., 2006b). Studies in pigs have confirmed the presence of AR in lymph (Marklund et al., 2014b). AR can be found in different blood fractions, mainly erythrocyte membranes and isolated lipoprotein fractions from plasma (very-low-density lipoprotein and high-density lipoprotein). AR concentrations in erythrocyte membranes and lipoprotein are highly correlated, and higher proportions of longer AR homologs are incorporated into erythrocyte membranes (Linko-Parvinen et al., 2007, Linko and Adlercreutz, 2005).

AR are also distributed in adipose tissue, as first indicated by radio activity measured in adipose tissue of rats fed a 14C-labeled C21:0 (Ross et al., 2004c) and later in a study where AR were measured in human adipose tissue biopsies (Jansson et al., 2010). However, the distribution of AR between blood and adipose tissue has not been studied in detail.

1.4.3 Elimination

Radioactive compounds detected in urine from rats fed 14C-labeled AR C21:0 (Ross et al., 2003b) have been identified as two main AR metabolites: 3,5-dihydroxybenzoic acid and 3-(3,5-dihydroxyphenyl)-1-propanoic acid (Ross et al., 2004a, Ross et al., 2004b, Ross et al., 2004d). A rat feeding experiment has also shown that co-ingestion of AR elevates the γ-tocopherol level in liver and lungs.
(Ross et al., 2004c), suggesting competition for metabolic enzymes, and thus, that AR might have similar metabolism to tocopherol (Birringer et al., 2001). Thus, it has been hypothesized that the alkyl chain in the AR molecule first undergoes ω-oxidation, followed by β-oxidation to form 3,5-dihydroxybenzoic acid and 3-(3,5-dihydroxyphenyl)-1-propanoic acid, which partly undergo phase II metabolism to form glucuronide and sulfate conjugates (Koskela et al., 2007). Several recently reported new metabolites of AR, including 2-(3,5-dihydroxybenzamido) acetic acid and 5-(3,5-dihydroxyphenyl) pentanoic acid and their glucuronide and sulfate conjugates, further support this hypothesis (Zhu et al., 2013). In addition, 3,5-dihydroxycinnamic acid has also been found to be a metabolite of AR from rye foods, but it might not come from AR through the route mentioned above (Bondia-Pons et al., 2013). AR metabolites are mainly excreted in urine (Marklund et al., 2013, Ross et al., 2004a, Ross et al., 2004b, Ross et al., 2004d), but some have also been found in bile from pigs and ileostomal effluent from humans (Marklund et al., 2014b).

The apparent elimination half-life of AR in plasma is estimated to be 4-5 h (Landberg et al., 2006b, Landberg et al., 2009a, Marklund et al., 2014b) whereas the half-life of the two main AR metabolites (3,5-dihydroxybenzoic acid and 3-(3,5-dihydroxyphenyl)-1-propanoic acid) in plasma was somewhat longer (10.1-16.3 h) (Söderholm et al., 2009) and in urine (9.9-13.5 h) (Söderholm et al., 2011). Thus, AR in plasma and their metabolites in plasma and urine may only reflect short-term whole grain intake, particularly in populations with low and irregular intake. More long-term biomarkers are therefore highly warranted.
1.5 Alkylresorcinols and their metabolites as dietary biomarkers

1.5.1 Dietary biomarkers

Biomarkers can be defined in several ways. The National Institutes of Health defines a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.” (Strimbu and Tavel, 2010). Biomarkers can be classified into three categories: exposure biomarkers, susceptibility biomarkers and effect biomarkers (Council, 1987). Exposure biomarkers should objectively reflect the xenobiotic material presented to an organism. Such a xenobiotic material can be a dietary component related to a specific food item. In nutritional studies, several biomarkers have been evaluated as indicators of food exposure and/or nutritional status (Bowman et al., 2011, Eisert et al., 2005, Lampe, 2003, Prentice et al., 2002, Yin et al., 2015). These biomarkers are referred to as dietary biomarkers, and can either represent the actual concentration or content of the nutrient or bioactive compound of interest, its metabolites, or its related metabolites in various biospecimen (Rappaport et al., 2014).

Biomarker classification

Dietary biomarkers can be classified as recovery biomarkers, concentration biomarkers, replacement biomarkers, and predictive biomarkers (Corella and Ordovas, 2015, Van Dam and Hunter, 2013).

Recovery biomarkers aim to provide unbiased estimates of the intake over a given period of time based on the concept of the balance of intake and excretion (recovery usually >80%) without being significantly influenced by differences in absorption, distribution, metabolism, and excretion (Bingham, 2002). Unfortunately, there are only a few biomarkers that qualify for this group, including doubly labeled water for average energy expenditure, and 24-h urinary nitrogen, sodium, and potassium for intake of protein, sodium, and potassium, respectively (Freedman et al., 2014, Kawasaki et al., 1993, Schoeller and van Santen, 1982, Schoeller et al., 1986).

Concentration biomarkers on another hand, reflect the concentration of target compound at a specific time point. Concentration biomarkers are affected not only by the dietary intake, but also by non-dietary determinants causing variations in absorption, distribution, metabolism, and excretion of the biomarker (Van Dam and Hunter, 2013). For example, vitamin C concentration...
in serum has been used as a biomarker of vegetable and fruit intake (Drewnowski et al., 1997). However, the concentration of vitamin C is also affected by the smoking status of participant (Schectman et al., 1989).

Some concentration biomarkers for compounds with a poor or even no intake database have been classified as replacement biomarkers. The best example of replacement biomarkers are phytoestrogen concentrations in serum and urine as dietary biomarker of phytoestrogen intake (Grace et al., 2004).

Predictive biomarkers, share commonalities with recovery biomarkers but are recovered to a far lower extent (Tasevska et al., 2005). For example, sucrose and fructose in 24-h urine reflect and predict sugar consumption in subjects with less influence by factors other than intake than for concentration biomarkers but with a low recovery (< 0.1% of ingested dose) (Tasevska et al., 2005).

Biomarker evaluation

Any suggested dietary biomarker must be thoroughly evaluated (Bearer et al., 2004). Validity and reproducibility of a dietary biomarker are two important features in determining biomarker performance. The validity of a dietary biomarker refers to its sensitivity and specificity in response to true dietary intake, i.e., the degree to which the biomarker measure what it is supposed to measure (Hunter et al., 2010, Kaaks, 1997). As there are no method that can measure diet without any measurement error (i.e., the true intake), repeated weighed food records have often been used as a proxy for true intake for biomarker validation (Eysteinsdottir et al., 2012, Kaaks, 1997). While comparing the biomarker with intake from FFQ, weaker correlations are typically observed as a result of relatively larger measurement error inherent in FFQ compared to food records. Correlations between an imperfect measure of dietary intake (due to measurement errors) and the biomarker reflect ‘relative’ validity (Masson et al., 2007, Ocke et al., 1997).

Since specific food intake varies over time, different numbers of repeated assessments are needed to capture the long-term mean intake of specific foods/nutrients (Fung et al., 2008). The reproducibility describes the correlation between repeated measurements within the same individual, i.e., the degree of fluctuation of the biomarker over time within an individual (Kaaks, 1997). This is often assessed by determination of the intra-class correlation coefficient (ICC) (Bartko, 1966), calculated as:

$$ICC = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_w^2}$$

Where $\sigma_B^2$ and $\sigma_w^2$ are inter- and intra-individual variation, respectively.
The ICC is an estimate of the proportion of total variation that is due to variation between individuals. In other words, a high ICC indicates a small variation within an individual in relation to the variation between individuals, and thus a single measurement of the dietary biomarker is sufficient to reflect long-term food intake, which is often of primary interest in epidemiological studies (Hankinson et al., 1995).

1.5.2 Alkylresorcinols in plasma and their metabolites in plasma or urine as intake biomarkers

AR concentrations in plasma and the concentrations of their main metabolites (3,5-dihydroxybenzoic acid and 3-(3,5-dihydroxyphenyl)-1-propanoic acid) in urine have been evaluated as concentration biomarkers in several intervention and observational studies (Andersson et al., 2011, Cuff et al., 2015, Landberg et al., 2009a, Marklund et al., 2013, McKeown et al., 2016a, McKeown et al., 2016b, Nybacka et al., 2016). The correlations between intake of whole grain or cereal fiber (as a surrogate) and plasma total AR concentrations were strongest (r=0.34-0.58) in studies using instruments with smaller measurement errors, e.g., weighed food records and FFQ specifically designed to capture whole grain intake, and weaker in studies with general FFQs (r=0.12-0.44) (Table 5).

Although the apparent half-life of plasma AR is relatively short, it reflected the whole grain intake and source of whole grain (wheat or rye) in populations with high and regular intake (Andersson et al., 2011, Landberg et al., 2013, Montonen et al., 2010). It has also been used in studies of different endpoints (Biskup et al., 2016, McKeown et al., 2016a, Olsen et al., 2010). Moreover, plasma AR concentration has been used to assess compliance in several whole grain intervention studies (Kristensen et al., 2012, Ross et al., 2011). A recent whole grain intervention study used plasma AR concentration to monitor the compliance of participants and observed that >60% of participants in the whole grain group did not follow the study design (Kristensen et al., 2017). Another study used AR concentrations along with other biomarkers as the basis to remove subjects with poor compliance in a secondary analysis of the outcome of a lifestyle intervention, and found significant improvements in the effects observed (Marklund et al., 2014a).

A main advantage of using intact plasma AR as biomarkers compared to their metabolites is the possibility to reflect total whole grain wheat and rye intake by measuring total AR concentration, but also the possibility to assess the relative intake of whole grain rye and wheat through measuring the AR C17:0/C21:0 ratio (Kristensen et al., 2012, Kyrø et al., 2014a, Linko-Parvinen et al., 2007, Magnusdottir et al., 2014a). The plasma AR C17:0/C21:0 ratio has been shown to be lowest (around 0.1) in the UK with mainly wheat as the source of whole
grain intake (Kyrø et al., 2014a, Mann et al., 2015, Ross et al., 2007), modest (around 0.3) in Sweden and Denmark, where both whole grain wheat and rye are consumed (Kyrø et al., 2013, Kyrø et al., 2014a, Ross et al., 2007), and highest (around 0.6) in Finland, with relatively high consumption of whole grain rye (Kyrø et al., 2014a, Montonen et al., 2003).

Table 5. Correlation of total plasma alkylresorcinol (AR) concentration with whole grain/cereal fiber intake from food frequency questionnaires (FFQ) and food records.

<table>
<thead>
<tr>
<th>Region</th>
<th>n</th>
<th>Intake variable</th>
<th>Method</th>
<th>Correlation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>28</td>
<td>Whole grain</td>
<td>3-d weighed food records</td>
<td>Spearman's r=0.58</td>
<td>(Landberg et al., 2008a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>51</td>
<td>Whole grain</td>
<td>3-d weighed food records</td>
<td>Spearman's r=0.38</td>
<td>(Andersson et al., 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>39</td>
<td>Rye bread</td>
<td>4-d food records</td>
<td>Spearman's r=0.34</td>
<td>(Linko et al., 2005)</td>
</tr>
<tr>
<td>Finland</td>
<td>56</td>
<td>Cereal fiber</td>
<td>3-d food records</td>
<td>Pearson's r=0.41</td>
<td>(Aubertin-Leheudre et al., 2008b)</td>
</tr>
<tr>
<td>Switzerland</td>
<td>29</td>
<td>Whole grain</td>
<td>3-d weighed food records</td>
<td>Pearson's r=0.57</td>
<td>(Ross et al., 2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>360</td>
<td>Rye bread</td>
<td>FFQ</td>
<td>Spearman's r=0.25</td>
<td>(Landberg et al., 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crisp bread</td>
<td>FFQ</td>
<td>Spearman's r=0.12</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>252</td>
<td>Whole grain</td>
<td>FFQ</td>
<td>Spearman's r=0.33-0.44</td>
<td>(Ross et al., 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>407</td>
<td>Whole grain</td>
<td>FFQ</td>
<td>Spearman's r=0.31</td>
<td>(Ma et al., 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wheat-rich foods</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Cereal-specific FFQ with 43-item on cereal products.

Beside AR concentrations in plasma, AR metabolites in plasma (Aubertin-Leheudre et al., 2009, Aubertin-Leheudre et al., 2010), spot urine (Wierzbicka et al., 2017), 12-h urine (Guyman et al., 2008), 24-h urine (Landberg et al., 2009a, Marklund et al., 2013) and 72-h urine have also been studied as biomarkers of whole grain intake (Aubertin-Leheudre et al., 2008, Aubertin-Leheudre et al., 2010). AR metabolites in these matrices show a similar correlation to whole grain intake as intact AR in plasma and similar reproducibility (Cuff et al., 2015), but slightly longer half-life (~10h).
However, AR metabolites cannot reflect the source of whole grain (wheat or rye).

### 1.5.3 Alkylresorcinols in adipose tissue as long-term biomarkers?

Since long-term dietary exposure is probably most relevant in studies of the role of whole grain in chronic disease prevention, biomarkers for long-term intake of separate whole grains are greatly needed, particularly for studies in populations with low and irregular intake of whole grain.

Lipophilic compounds, such as fatty acids, have a longer half-life in adipose tissue than in blood (Katan et al., 1997, Strawford et al., 2004) and have been used as long-term biomarkers of diet in several studies (Ito et al., 1991, Petrek et al., 1997, Schmidt et al., 2014). Inspired by this concept, total AR concentration in adipose tissue has been measured and shown to be well correlated ($r=0.48$) with long-term whole grain bread intake from FFQ in a small pilot study (Jansson et al., 2010). This suggested AR in adipose tissue has large potential to be used as long-term biomarker of whole grain wheat and rye intake as well as their relative intakes. Such biomarkers could be used to investigate the potential role of whole grain rye for cholesterol reduction as well as investigating the controversial role of whole grains, particularly rye, in prevention of breast cancer. However, the analytical method for AR in adipose tissue reported in the pilot study (Jansson et al., 2010) does not allow analysis with high precision nor high-throughput. Moreover, studies to evaluate AR in adipose tissue as long-term biomarkers by investigating response to controlled intakes, correlations to estimated intakes by different dietary assessment methods and in different study settings were lacking before the initiation of the research in this thesis.
2 Aims of thesis

The overall aims of this thesis were to evaluate if AR in adipose tissue could be used as long-term biomarkers of whole grain wheat and rye intake and to investigate whether these biomarkers, in parallel with whole grain intakes determined by food frequency questionnaires, are associated with blood lipid profiles and breast cancer incidence in the Nordic populations. The aims were addressed through the following objectives:

1) Develop a robust, high-throughput, sensitive and affordable GC-MS method for determination of AR concentrations in a small volume (10-50mg) of adipose tissue biopsy sample (Paper I).

2) Investigate the response of AR concentrations in adipose tissue during controlled dietary intervention conditions using samples and data from a 12-week whole grain intervention study (Paper II).

3) Examine potential determinants of AR concentrations in adipose tissue among free-living men and women (Paper III).

4) Compare the performance of AR concentrations in adipose tissue and in plasma as biomarkers of whole grain wheat and rye intake (Papers II-III).

5) Investigate whether long-term intake of whole grain rye and/or wheat are associated with favorable blood lipid profiles, using FFQ-based estimates of whole grain intake, total AR concentrations and AR C17:0/C21:0 ratio in adipose tissue and plasma as exposure measurements (Paper IV).

6) Investigate whether AR in adipose tissue, as biomarkers of whole grain rye and/or wheat, are associated with risk of developing breast cancer in a case-subcohort study in Danish women (Paper V).
3 Materials and methods

3.1 Reference compounds
Analyzed alkylresorcinols (C17:0, C19:0, C21:0, C23:0 and C25:0) and AR used as internal standards (C20:0, C22:0, C24:0, C26:0) were of >95% purity and were provided by Researchem Lifescience (Burgdorf, Switzerland). Similar weight of analyzed AR homologs were solubilized in methanol and diluted with methanol to nine different concentrations that reflected the range of AR likely to be found in plasma and in adipose tissue biopsy samples (Table 6). Alkylresorcinol internal standards (C20:0, C22:0, C24:0, C26:0) were solubilized in methanol (300 µg/L for each internal standard AR homologue) and used for preparation of calibration curves and in samples.

3.2 Instrumentation
AR in samples were quantified using two gas chromatography-single quadrupole mass spectrometry (GC-MS) instruments (Finnigan Trace GC/Finnigan Trace DSQ II, Thermo Fisher Scientific, Waltham, MA, USA). A stainless steel liner was installed in the spilt/splitless inlet on one of the GC-MS, while a straight glass liner was installed in the other GC-MS. ZB-5MS columns (15m×0.25mm×0.25µm, Zebron) were installed on both instruments. Helium (1.0mL/min) was used as carrier gas.

3.3 Alkylresorcinols analysis in adipose tissue and plasma
In the study described in Paper I, a method for quantification of AR in adipose tissue biopsies was developed and evaluated based on a method for AR measurement in plasma (Landberg et al., 2009b). The method comprises four main steps: (1) Sample extraction; (2) Solid phase extraction (SPE) clean-up; (3) Derivatization, and (4) GC-MS analysis.
Table 6. Standard solutions for calibration (n=9) and their corresponding target alkylresorcinol (AR) concentration in plasma or content in adipose tissue.

<table>
<thead>
<tr>
<th>Standard mixture</th>
<th>C17:0 nmol/L(^1) pmol/g(^2)</th>
<th>C19:0 nmol/L(^1) pmol/g(^2)</th>
<th>C21:0 nmol/L(^1) pmol/g(^2)</th>
<th>C23:0 nmol/L(^1) pmol/g(^2)</th>
<th>C25:0 nmol/L(^1) pmol/g(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.5 14.0</td>
<td>2.5 10.0</td>
<td>2.8 11.2</td>
<td>2.3 9.2</td>
<td>2.2 8.8</td>
</tr>
<tr>
<td>2</td>
<td>7.0 28.0</td>
<td>5.0 20.0</td>
<td>5.5 22.0</td>
<td>4.5 18.0</td>
<td>4.4 17.6</td>
</tr>
<tr>
<td>3</td>
<td>20.9 83.6</td>
<td>15.0 60.0</td>
<td>16.6 66.4</td>
<td>13.6 54.4</td>
<td>13.2 52.8</td>
</tr>
<tr>
<td>4</td>
<td>34.8 139.2</td>
<td>25.0 100.0</td>
<td>27.7 110.8</td>
<td>22.6 90.4</td>
<td>22 88.0</td>
</tr>
<tr>
<td>5</td>
<td>69.5 278.0</td>
<td>50.0 200.0</td>
<td>55.4 221.6</td>
<td>45.3 181.2</td>
<td>43.9 175.6</td>
</tr>
<tr>
<td>6</td>
<td>104.3 417.2</td>
<td>75.0 300.0</td>
<td>83.2 332.8</td>
<td>67.9 271.6</td>
<td>65.9 263.6</td>
</tr>
<tr>
<td>7</td>
<td>208.6 834.4</td>
<td>150.0 600.0</td>
<td>166.3 665.2</td>
<td>135.8 543.2</td>
<td>131.7 526.8</td>
</tr>
<tr>
<td>8</td>
<td>347.7 1390.8</td>
<td>250.0 1000.0</td>
<td>277.2 1108.8</td>
<td>226.3 905.2</td>
<td>219.6 878.4</td>
</tr>
<tr>
<td>9</td>
<td>417.2 1668.8</td>
<td>300.0 1200.0</td>
<td>332.7 1330.8</td>
<td>271.5 1086.0</td>
<td>263.5 1054.0</td>
</tr>
</tbody>
</table>

\(^1\) Corresponding AR concentration in plasma if the standard mixture would have been spiked to 200 µL blank plasma.

\(^2\) Corresponding AR concentration in adipose tissue sample if the standard mixture would have been spiked to 50 mg blank adipose tissue sample.
3.3.1 Sample extraction

Plasma sample extraction
Plasma (200 µL) was vortexed with internal standard (18.0 ng of AR in 15 µL methanol, which contained equal weight of AR C20:0, AR C22:0, AR C24:0, AR C26:0), and extracted with diethyl ether (3 mL), for a total of three times. The organic extract was combined and dried at 35 °C under a gentle stream of nitrogen in a glass test tube. Extracted AR in the dried glass tube were reconstituted in methanol (1 mL), ready for SPE clean-up (Landberg et al., 2009b).

Adipose tissue extraction
Adipose tissue (10-50 mg) was homogenized using a bead homogenizer with internal standard solution (18.0 ng of AR in 15 µL methanol, which contained equal weight of AR C20:0, AR C22:0, AR C24:0, AR C26:0) and organic solvent A was added (1 mL of diethyl ether, methanol, or acetone was tested). The organic extract was transferred to a new tube and evaporated to dryness at 35 °C. The dried extract was then further extracted with organic solvent B (1 mL of diethyl ether, methanol or acetone was tested). Organic extract from the second extraction were either cooled at 4 °C for 30 min followed by filtration with a 0.45 µm syringe filter, or centrifuged (0 °C, 20 817× g, 10 min). The clear extract/supernatant was ready for SPE clean-up.

3.3.2 Solid phase extraction clean-up
SPE-cleanup of plasma and adipose tissue was performed using either a four-channel SPE robot (flow rate:1mL/min) with Oasis® MAX 60 mg SPE cartridges (Waters, Milford, MA, USA) or an 8-channel pipette with Oasis® MAX 96-well 60 mg SPE plate (under 5 mm Hg).

The MAX cartridges were activated with 1mL of conditioning solution (0.1 M NaOH : methanol, 3:7 v/v), loaded with the clear sample extract, washed with methanol (3 mL), and finally eluted with 3 mL 2% formic acid in methanol (Figure 2). The eluted samples were evaporated to dryness under a nitrogen stream at 60 °C.

The same protocol as for MAX cartridges was applied to the MAX 96-well SPE plate, except that 4 mL instead of 3mL of 2% formic acid in methanol was used to elute AR from the plate during method development.
Figure 2. Sample cleanup with Oasis® MAX solid phase extraction cartridges. The MAX cartridges/plates were first activated with 1 mL of conditioning solution (0.1 M NaOH: methanol, 3:7 v/v), and then sample contained unwanted matrix (yellow droplet) was loaded (A) and then cartridge was washed with 3 mL methanol (B). Unwanted matrix (yellow droplet) was eluted and discarded, while alkylresorcinols were retained on cartridge. Alkylresorcinols were eluted from the cartridge with 3 mL (or 4 mL during method development) 2% formic acid in methanol (C).

3.3.3 Derivatization

AR from adipose tissue were derivatized with trifluoroacetic anhydride (TFAA), while for AR in plasma, TFAA was compared with quick silylation mixture (QSM, pyridine:hexamethyldisilazane:trimethylchlorosilane, 9:3:1 v/v/v). Reaction conditions for the TFAA derivatization, including reaction time (30-60 min), temperature (30-60 °C), reconstitution method (vortexing and/or heating), and stability of derivatized samples were investigated.

3.3.4 Analysis of alkylresorcinols on GC-MS

Plasma samples derivatized with QSM were analyzed using the method previously described (Landberg et al., 2009b). Briefly, sample (1.5 μL) was injected into the GC-MS fitted with a straight glass liner (300 °C) in splitless mode. The oven temperature was held at 100°C (0-1 min), raised to 250 °C by
8.6 °C/min, then raised to 300°C by 40°C/min, and held at 300°C for 2 min. The transfer-line and ion source was held at 310 °C and 250°C, respectively. Selected ion monitoring (SIM) under electron-impact ionization (EI) at 70eV was used for quantification, with a quantification ion (in bold) and a confirmation ion for each AR homolog: m/z 492 and 268 for C17:0, m/z 520 and 268 for C19:0, m/z 548 and 268 for C21:0, m/z 576 and 268 for C23:0, m/z 604 and 268 for C25:0, and m/z 520 and 268 for C20:0 (internal standard).

Plasma and adipose tissue samples derivatized with TFAA were injected into the GC-MS fitted with a stainless steel liner. The transfer-line and ion source was held at 310 °C and 250°C, respectively. The GC conditions, including inlet temperature (250-320 °C), and oven temperature were optimized. The EI (70eV) and negative-chemical-ionization (NCI) modes were tested and compared.

3.3.5 Calibration

Standard solutions (Table 6) were randomly analyzed along with all samples in each batch. Known AR concentrations were linearly regressed against the ratio of target-AR/ internal-standard-AR. AR C20:0 was used to calibrate all target AR (C17:0-C25:0) in plasma samples, while different internal standard AR homologs were used for different AR homologs in adipose tissue: C20:0 for C17:0 and C19:0, C22:0 for C21:0, and C24:0 for C23:0 and C25:0.

3.3.6 Evaluation of the method

The linearity of the GC-MS was evaluated using standard solutions (Table 6) which covered likely AR concentrations in the needle biopsy adipose samples (usually around 10-50 mg). The linearity of the method was tested using six adipose tissue samples (10.2-54.0 mg) from one pig homogenized with a random number of stainless steel beads (n=1-6). The theoretical limit of detection (LOD) and limit of quantification (LOQ) were taken as the AR concentrations where the signal to noise ratio was 3 and 10, respectively. The precision of the method was estimated by analyzing AR concentration in replicates of a quality control adipose tissue (n=4 per day, repeated 3 times during one week). The accuracy of the method was evaluated by spiking the quality control adipose tissue with three levels of AR mixtures (10 ng, 20 ng and 30 ng of total AR; equal mass for all target AR homologs).
3.4 Samples and study designs

Populations and study design used for the human studies reported in Papers II - V are summarized in Table 7 and short descriptions are provided below.

3.4.1 Populations, subjects and samples

**SYSDIMET (Paper II)**

Data, plasma and adipose tissue biopsy samples were obtained from the Dietary Modulation of Gene Expression and Metabolic Pathways in Glucose Metabolism (SYSDIMET, ClinicalTrials.gov Identifier: NCT00573781) study conducted on Finnish participants (n=131) (de Mello et al., 2011). All participants were aged 47-65 years, had impaired glucose metabolism (fasting plasma glucose concentration in the range 5.2-7.8 mmol/L and 2-h plasma glucose concentration in the range 7.8-11.0 mmol/L), and met at least two of the following criteria: body mass index (BMI) 26-39 kg/m²; waist circumference ≥ 88 cm in women and ≥ 102 cm in men; blood pressure ≥ 130/≥ 85 mm Hg or use of medication for hypertension; serum triglycerides ≥ 1.7 mmol/L; and high-density lipoprotein-cholesterol (HDL-C) <1.3 mmol/L in women and <1.0 mmol/L in men at baseline. During the 12-week intervention, all participants were randomly allocated to one of three dietary intervention groups: (1) A refined diet group (RDG) (n=45) consuming refined wheat products; (2) a healthy diet group (n=44) with high intake of whole grain, fatty fish, and blueberries; and (3) a whole grain-enriched diet group (n=42) with high intake of whole grain. Participants in the healthy diet group and the whole grain-enriched diet group were asked to replace their other habitual cereal products with products containing more than 50% of whole grain and to consume more than 35 g (weight before cooking) whole grain pasta per day. In addition, a daily portion of their habitual cereal product was permitted, and oat snack bars were provided (one a day on a voluntary basis).

Since the whole grain intake guideline was the same for the healthy diet group and the whole grain-enriched diet group, these two groups were merged to one whole grain diet group (WGDG) for the purpose of the study. In total, plasma and adipose tissue samples from 17 subjects in the WGDG group and 10 subjects in the RDG group, collected at baseline and at end of the study (week 12), were available for analysis.
Table 7. Design of human studies (Papers II- V) \(^1\) (AR = alkylresorcinols, FFQs = food frequency questionnaires)

<table>
<thead>
<tr>
<th>Paper</th>
<th>Main objective</th>
<th>Country</th>
<th>Subjects</th>
<th>Age(^2) (years)</th>
<th>BMI(^2) (kg/m(^2))</th>
<th>Type of study</th>
<th>Self-reported method</th>
<th>Biopsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Evaluate AR as biomarkers</td>
<td>Finland</td>
<td>27</td>
<td></td>
<td>12 wk intervention with two treatment groups. Whole grain diet group(^3) (WGDG, n=17) and refined diet group (n=10)</td>
<td>Whole grain intake estimated by 4 day food records before the intervention (week 0) and in week3, 7, and 11 of the intervention.</td>
<td>Adipose and plasma collected at baseline and after 12-wk intervention</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Evaluate AR as biomarkers</td>
<td>Sweden</td>
<td>149 men</td>
<td>87 (85, 88)</td>
<td>24.6 (22.6, 27.2)</td>
<td>Cross-sectional study. Men were randomly selected from the Cohort of Swedish Men-Clinical; women were randomly selected from the Swedish Mammography Cohort-Clinical</td>
<td>Men: whole grain intakes were reported using repeated FFQs in 1997, 2009 and 2010.</td>
<td>Adipose tissue and plasma collected in 2010</td>
</tr>
<tr>
<td></td>
<td>Evaluate AR as biomarkers</td>
<td>Sweden</td>
<td>109 women</td>
<td>66 (60,71)</td>
<td>24.3 (23.0, 25.7)</td>
<td>Cross-sectional study. Women were randomly selected from the Swedish Mammography Cohort-Clinical</td>
<td>Women: whole grain intakes were reported using repeated FFQs in 1987, 1997 and 2003.</td>
<td>Adipose tissue and plasma collected in 2003</td>
</tr>
<tr>
<td>IV</td>
<td>Associate whole grain and blood lipids</td>
<td>Sweden</td>
<td>109 women</td>
<td>66 (60,71)</td>
<td>24.3 (23.0, 25.7)</td>
<td>Cross-sectional study. Women were randomly selected from the Swedish Mammography Cohort-Clinical</td>
<td>Whole grain intakes were reported using by FFQ in 2003.</td>
<td>Adipose tissue and plasma collected in 2003</td>
</tr>
<tr>
<td>V</td>
<td>Associate whole grain and breast cancer risk</td>
<td>Denmark</td>
<td>1 347 women</td>
<td>61 (58, 65)</td>
<td>24.6 (22.5, 27.5)</td>
<td>Case-subcohort study between 01/01/2001 and 27/04/2006</td>
<td>Whole grain intakes were reported using FFQ distributed once at entry (1993 and 1997).</td>
<td>Adipose tissue collected in a single sample at entry (1993-1997)</td>
</tr>
</tbody>
</table>

\(^1\) Women in Papers III and IV are the same group of people.
\(^2\) Median (25\(^{th}\) percentile, 75\(^{th}\) percentile).
\(^3\) Participants from the healthy diet group and the whole grain-enriched diet group in SYSDIMET (de Mello et al., 2011) consumed similar amounts of whole grain, and were therefore merged to one whole grain diet group (WGDG) in Paper II.
Swedish Mammography Cohort and the Cohort of Swedish Men (Paper III and IV)

Between 1987 and 1990, the Swedish Mammography Cohort (SMC) (Figure 3) was initiated by inviting 90,303 women, born between 1914 and 1948 and living in the counties of Västmanland and Uppsala, central Sweden, by sending a questionnaire concerning diet (67-food items), height, weight, education, and marital status (FFQ-87). The response rate was 74%. In 1997, all SMC participants (n=56,030) who were still living in Västmanland and Uppsala were mailed a more comprehensive questionnaire which also included questions on dietary supplements, physical activity and smoking status (FFQ-97) (Wolk et al., 2006), and response rate was 70%. Between 2003 and 2009, SMC women (n=8,311) living in Uppsala County were invited to join SMC-Clinical. These women were asked to fill out a 123-item FFQ (FFQ-03); undergo a health examination, including dual energy X-ray absorptiometry scan, registration of body weight, height, blood pressure, and waist and hip circumference; and adipose tissue biopsies, fasting blood- and urine samples were collected. The response rate was 60%.

The Cohort of Swedish Men (COSM) (Larsson et al., 2005) was initiated in 1997 by inviting men living in the counties of Västmanland and Örebro, central Sweden and born between 1918 and 1952 (Figure 3). A total of 48,850 men filled out a questionnaire (as the FFQ-97 in the SMC). One more food related questionnaire (FFQ-09) was sent to all COSM members (n=37,861) who still lived in the study area in 2009, to which the response rate was 78%. Between 2003 and 2009, SMC women (n=8,311) living in Uppsala County were invited to join SMC-Clinical. These women were asked to fill out a 123-item FFQ (FFQ-03); undergo a health examination, including dual energy X-ray absorptiometry scan, registration of body weight, height, blood pressure, and waist and hip circumference; and adipose tissue biopsies, fasting blood- and urine samples were collected. The response rate was 60%.

In total, data, adipose tissue biopsies, and plasma samples were available for 109 SMC-C and 149 COSM-C participants in Papers III and IV, respectively. Blood lipids of SMC-Clinical participants had been measured in a previous study and the data were available.

Danish Diet, Cancer and Health Study (Paper V)

The Danish “Diet, Cancer and Health” cohort (Tjønneland et al., 2007) was started between 1993 and 1997, by inviting 80,996 men and 79,729 women, aged 50-64, born in Denmark, living in Copenhagen and Aarhus area and were not
Figure 3. The Swedish Mammography Cohort (SMC) between 1987 and 2009, and the Cohort of Swedish Men (COSM) between 1997 and 2013. Women (n=109) were randomly selected from the SMC Clinical sub-study responders for Paper III and IV; men (n=149) were randomly selected from the COSM Clinical sub-study responders for Paper III.
registered with a previous diagnosis of cancer in the Danish Cancer Registry (Gjerstorff, 2011). A total number of 27 178 men and 29 875 women joined the cohort between 1993 and 1997. All participants were asked to fill out a 192-item FFQ prior to visiting one of two study centers. At the study centers, participants were asked to complete a lifestyle questionnaire concerning years in school, use of HRT, smoking status, menopause status, parity, age at menarche, age at first birth, age at menopause, and physical activity. In addition, anthropometric measurements, including weight, height, sitting height, and waist and hip circumference were performed. A blood sample, a spot urine sample and toenail clippings were collected from each participant. A needle biopsy of adipose tissue (35-50 mg) was taken from the buttock, frozen (-20 °C) within 2 hours of collection, and stored in liquid vapor (max.-150 °C). Vital status, migration, and disease development were followed up continuously. Breast cancer cases during the follow-up were identified by the Danish Breast Cancer Cooperative Group, the Danish Cancer Registry (Gjerstorff, 2011), and the Danish Pathology Register (Bjerregaard and Larsen, 2011).

Data and adipose tissue from women who were diagnosed with breast cancer between date of entry and December 2000 have been used in a previous study (Raaschou-Nielsen et al., 2005). A case-subcohort study (Witt et al., 2009) was conducted by following women for breast cancer from January 1, 2001 (new entry point) until date of cancer diagnosis, date of death, date of emigration, or April 27, 2006 (chosen as end of follow-up), which ever came first. Adipose tissue samples from this sub-cohort study were available (Witt et al., 2009), and therefore the same study design was used with a few modifications. In brief, all invasive breast cancer cases (n=500) diagnosed between January 1, 2001 and April 27, 2006 and a sub-cohort of non-cases (n=1115) randomly selected from the entire cohort were included. After excluding participants with missing adipose tissue biopsies, a case group (n=414; 345 ER+, 56 ER-, and 13 ER status unknown) and a subcohort of 933 non-cases (none of whom developed breast cancer during follow-up) remained and were included in the analysis.

3.4.2 Dietary assessment of whole grain intake

SYSDIMET (Paper II)

Four-day food records with consecutive pre-defined days including one weekend day were used to estimate food intake in week before the intervention (week 0) and during weeks 3, 7, and 11 of the intervention. Food and nutrients were calculated using Nutrica software (https://fineli.fi/fineli/fi). Whole grain intake during weeks 3, 7, and 11 were averaged as whole grain intake during the
intervention. Whole grain intake information was missing for 11 participants in the WGDG.

**SMC and COSM (Paper III and IV)**

FFQs (FFQ-87, FFQ-97, and FFQ-03 for SMC; FFQ-97, FFQ-09, and FFQ-10 for COSM) (Figure 3) were used to collect data on average food intake frequencies during the previous year (Thomas et al., 2011, Wallin et al., 2014, Warenstjö et al., 2011). Frequencies were converted to average food intake (g/day) by multiplying age-specific portion sizes (based on data from weighed food records in 159 men over 2 weeks and 213 women over 4 weeks distributed over one year). Average intake frequency of whole grain foods, including ‘crispbread’, ‘white bread’, ‘whole-meal bread’, ‘porridge’, ‘muesli’, ‘pasta’, and ‘pancake’, was multiplied by the specific average content of total whole grain, whole grain rye, and whole grain wheat listed for the corresponding food categories in the Swedish national food composition database (Bergström et al., 1991). The average daily intake (g/day), including total whole grain intake (TWG), whole grain rye intake (WGR), and whole grain wheat intake (WGW), for every FFQ was calculated. Total whole grain rye and wheat intake (WGR&W) was calculated as the sum of WGW and WGR. Relative whole grain rye intake (WGR%) was calculated as the ratio of WGR to WGR&W. Average whole grain intake from FFQs during years prior to adipose/plasma biopsy was calculated: for men as the mean of all three FFQs (FFQ-97, FFQ-09, and FFQ-10) for 14 years prior to adipose tissue and plasma biopsy and the mean of FFQ-09 and FFQ-10 for two years prior to biopsy; and for women as the mean of all three FFQs (FFQ-87, FFQ-97, and FFQ-03) for 17 years prior to biopsy, and the mean of FFQ-97 and FFQ-03 for 7 years prior to biopsy.

**Danish Diet, Cancer and Health Study (Paper V)**

FFQ-based estimates, including mean daily intake of total energy, whole grain rye, and whole grain wheat, were calculated using the approach described in previous studies (Egeberg et al., 2009, Haraldsdottir et al., 1994, Helnæs et al., 2016, Kyrø et al., 2011, Overvad et al., 1991, Tjønneland et al., 1991, Tjønneland et al., 1992). Total whole grain rye and wheat intake (WGR&W) was calculated as the sum of WGW and WGR. Relative whole grain rye intake (WGR%) was calculated as the ratio of WGR and WGR&W.
3.5 Statistical analysis

All statistical analyses were conducted using SAS 9.3/9.4 for Windows (SAS Institute Inc., Cary, NC, USA). Two-sided $P<0.05$ was defined as statistically significant. Methods and models used in different studies are summarized in Table 8 and briefly described below.

**Paper II**

Whole grain intake and AR concentrations in adipose tissue and plasma were tested either with Wilcoxon’s signed-rank test between week 0 and week 12 for each group, or with Mann-Whitney’s U-test between two treatment groups at week 0 and week 12. Correlations between whole grain intake and AR concentrations in adipose tissue or in plasma were assessed by Spearman’s rank correlation coefficient.

**Paper III**

Concentration biomarkers were found to be right-skewed and were therefore log-transformed before statistical analysis. Spearman’s rank correlation coefficient was calculated for all correlations assessed. Analysis of variance (ANOVA) and ICC (calculated with %icc9 SAS macro) (Hertzmark and Spiegelman, 2010) were used to examine the differences between FFQ estimates and the reproducibility of FFQ-based whole grain intakes across repeated FFQs for both genders. Associations between energy-adjusted (residual method) whole grain intake (WGR and WGW) and total AR concentrations in adipose tissue or in plasma were further tested using two multiple linear regression models. In Model 1, total energy intake and energy-adjusted WGW and WGR were set as predictors, while total AR concentration in adipose tissue or in plasma (log-transform) was used as the outcome in separate models. In Model 2, sex (man or woman), age (as continuous variable), and BMI (as continuous variable) were added to Model 1 as covariates.
Table 8. Statistical methods, statistical procedures used in SAS, dependent and independent variables, and covariates used (AR=alkylresorcinols, FFQ=food frequency questionnaire).

<table>
<thead>
<tr>
<th>Paper</th>
<th>Statistical Procedure (SAS)</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>NPAR1WAY CORR</td>
<td>AR concentrations in adipose tissue and plasma at week 0 and after the 12-wk intervention</td>
<td>Whole grain intake at week 0 and during the 12-wk intervention</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>CORR %ICC MEANS ANOVA GLM</td>
<td>Women: AR concentration in adipose tissue and plasma collected in 2003</td>
<td>Women: Whole grain intake from FFQ-87, FFQ-97 and FFQ-03</td>
<td>Age, BMI&lt;sup&gt;1&lt;/sup&gt;, and sex</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Men: AR concentration in adipose tissue and plasma collected in 2010</td>
<td>Men: Whole grain intake from FFQ-97, FFQ-09 and FFQ-10</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>MEANS CORR GLM</td>
<td>Blood lipids measured in 2003</td>
<td>Whole grain intake from FFQ-03, total AR concentration and AR C17:0/C21:0 ratio in adipose tissue and plasma collected in 2003</td>
<td>Age and BMI&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>V</td>
<td>MEANS FREQ CORR PHREG</td>
<td>Breast cancer incidence</td>
<td>Whole grain intake from FFQ at entry, total AR concentration and AR C17:0/C21:0 ratio in adipose and collected at entry</td>
<td>BMI&lt;sup&gt;2&lt;/sup&gt;, education, parity, HRT&lt;sup&gt;2&lt;/sup&gt;, physical activity, smoking, alcohol, menopause, age at first birth, age at first menstruation.</td>
</tr>
</tbody>
</table>

<sup>1</sup> Body mass index.

<sup>2</sup> Hormone replacement therapy.
**Paper IV**

Total whole grain rye and wheat intake, and relative whole grain rye intake among female participants were assessed using three approaches (FFQ, AR in adipose tissue, and AR in plasma) in relation to blood lipids using three multiple linear models: 1) An FFQ-model where energy-adjusted WGR&W and WGR% were used as predictors; 2) an adipose-AR-model where total AR concentrations and AR17:0/C21:0 ratio in adipose tissue were used as predictors; and 3) a plasma-AR-model where total AR concentration and AR17:0/C21:0 in plasma were used as predictors. Blood lipids (TCHOL, LDL-C, HDL-C, and triglyceride) were log-transformed to improve normality and separately entered as dependent variables. Least-squares means and 95% confidence intervals of blood lipids were calculated across exposure tertiles using PROC GLM. Predictors were centered with the median of corresponding tertiles and entered into the models as continuous variables to calculate \( P \)-values for trend. Models were further adjusted for age and BMI (both entered as continuous variables).

**Paper V**

Spearman’s correlation coefficients between WGR&W and total AR concentration in adipose tissue, and between WGR% and AR C17:0/C21:0 ratio in adipose tissue, were calculated. The time-to-event hazard ratios (HR) of total breast cancer and of ER+ and ER- breast cancer were calculated using Cox’s partial likelihood (Cox, 1992). Age (in days) was used as the time scale in the models. Since the study followed a case-cohort design, the risk set for cases outside the subcohort were set to half a day before their time of event (Prentice, 1986). Total whole grain rye and wheat intake, and relative whole grain rye intake were estimated using FFQ or AR in adipose tissue, and used as exposures.

The following statistical models were evaluated:

1) **Crude FFQ-model**: Energy-adjusted WGR&W + WGR% + total energy intake.

2) **Crude AR-model**: Total AR concentration + AR C17:0/C21:0 ratio.

Exposures in the FFQ-model and the AR-model were either categorized according to quartiles or specified as continuous variables, where a one-unit increase represented: 10g/day of energy-adjusted WGR&W; 10% of WGR%; 1 nmol/g of total AR concentration in adipose tissue; and 0.1 of AR C17:0/C21:0 ratio in adipose tissue. The crude models were further adjusted for previously studied breast cancer risk factors, including BMI, education, parity, HRT, physical activity, smoking, alcohol, menopause status, and age at first birth/first period.
4 Results and discussion

4.1 Determination of alkylresorcinols in adipose tissue

The final method developed and evaluated in Paper I and used for analysis of AR in adipose tissue and plasma in Paper III-V is shown in Figure 4. The results from method development and evaluation are summarized and discussed below.

4.1.1 Sample extraction

For sample extraction, a bead mill homogenizer was compared with Heidolph Diax 600 homogenizer used in a previous method (Jansson et al., 2010). With the bead mill homogenizer, up to 24 samples could be analyzed at a time and potential cross-contamination between samples could be avoided. Different combinations of solvents A and B were tested for extraction and the combination of diethyl ether (A) and methanol (B) was selected, because all other combinations resulted in co-elution of large amounts of yellowish, oil-like impurities (even after further SPE clean-up), which caused poor baseline and required frequent maintenance of the GC-MS instruments. To remove remaining oil-like impurities extracted with the selected solvents, the sample extract was kept at 4 ºC for 30 min and impurities were crystalized and could be efficiently removed by filtration or centrifugation. Centrifugation was selected for the sample analysis in Paper V since it is cheaper than the filtration method and more robust.

4.1.2 Solid phase extraction clean-up

To increase throughput, an Oasis MAX 96-well plate with 60 mg sorbent was tested and evaluated. Most of the applied AR (82.7-94.8%) were eluted from the 96-well SPE plate during the first 1 mL of elution with 2% formic acid in methanol. Small amounts (<10%) of AR were found the second 1 mL of elution and the percentage was slightly higher for longer AR homologs, likely due to higher retention of longer AR homologues since Oasis MAX also consists of hydrophobic interaction. Only small amounts of C17:0 (12.7%) and C21:0 (5.5%) were lost during the washing step. In the final method, 3 mL of methanol was used during the SPE washing step to ensure the removal of matrix interfering compounds, while 3 mL of 2% formic acid in methanol elution were applied to the plate to ensure quantitative extraction of the AR applied. With the 96-well SPE plate, 92 unknown samples + 4 quality control samples could be purified.
Figure 4. Final protocol for plasma and adipose tissue extraction, solid phase extraction, derivatization and GC-MS analysis. Steps or consumables underlined were introduced after the publication of Paper I.
Figure 5. Total peak area percentage of alkylresorcinols homologs (C17:0-C25:0)) eluted from Oasis MAX solid phase extraction 96-well plate in three washing fractions using 1 mL methanol (ME01-ME03), followed by 4×1 mL elution fractions with 2% formic acid in methanol (FM01-FM04).

within two hours, which is 3× faster than the method with an SPE robot used previously (Landberg et al., 2009b).

4.1.3 Derivatization
Derivatization of AR with TFAA was conducted at 40 °C. Excess TFAA could damage the GC column and metal liner, and was therefore removed after derivatization by evaporation at 60 °C. Reaction times longer than 30 min did not improve the yield of derivatized AR and therefore 30 min was selected. The GC-vial containing TFAA-derivatized AR in undecane were heated at 60 °C for 10 min and vortexed for 20s to ensure that AR compounds were dissolved, since the viscosity of undecane is relatively high and some remaining oil-like compounds in the sample after the SPE clean-up could interfere with the reconstitution of AR in undecane. This heat-vortex treatment significantly increased the precision of the analysis (Table 9). AR concentrations of the same batch of samples injected 2 week apart showed excellent agreement ($r>0.99$, $P<0.001$ for all target AR homologues). This shows good stability of the generated AR derivatized in undecane and samples could be stored for up to at least two weeks before analysis.
Table 9. Alkylresorcinols (AR) concentrations (nmol/g) determined in a quality control adipose tissue sample with and without heat-vortex treatment (60°C for 10min and vortex for 20s) after derivatization.

<table>
<thead>
<tr>
<th>AR</th>
<th>Without heating (n=4)</th>
<th>With heating (n=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>C17:0</td>
<td>0.04</td>
<td>0.009</td>
</tr>
<tr>
<td>C19:0</td>
<td>0.58</td>
<td>0.075</td>
</tr>
<tr>
<td>C21:0</td>
<td>1.20</td>
<td>0.103</td>
</tr>
<tr>
<td>C23:0</td>
<td>0.12</td>
<td>0.025</td>
</tr>
<tr>
<td>C25:0</td>
<td>0.08</td>
<td>0.022</td>
</tr>
<tr>
<td>Total</td>
<td>2.03</td>
<td>0.205</td>
</tr>
</tbody>
</table>

4.1.4 GC-MS analysis and method evaluation

Injection volume (2 µL) was a compromise between sufficient sensitivity, ensuring low contamination in the liner to keep longer maintenance intervals and possibilities for reinjection of samples. The chromatographic settings enabled good separation of all AR homologs in the standard mixture (Figure 5) and in adipose tissue and plasma samples. The TFAA-derivatized AR were scanned under EI (70eV) and NCI mode. The EI mode was selected since comparable sensitivity was found for EI and NCI, while the throughput was much higher for EI than for NCI due to more frequent instrument maintenance when the instrument was operated in this mode (data not shown).

In EI mode, molecular ions and a common base ion (m/z 316) found in all TFAA-derivatized AR homologs were detected (Figure 6). Compared with molecular ions, the base ion had a higher signal to noise ratio, indicating good linearity (r²>0.99) across the concentrations of standard solutions. Based on these findings, the base-ion was selected for quantification with SIM mode. The theoretic LOD/LOQ per injection for AR were 0.04/0.12 fmol for C17:0, 0.05/0.16 fmol for C19:0, 0.02/0.08 fmol for C21:0, 0.02/0.07 fmol for C23:0, and 0.09/0.31 fmol for C25:0. Inter- and intra- batch coefficient of variation (CV) and accuracy for plasma and adipose tissue samples were satisfactory for most homologues at different concentration ranges (Table 10).
Figure 6. Comparison between Scan (A) and selective ion monitoring (SIM) modes (A-H) of an alkylresorcinol mixture (C17:0-C25:0) analyzed in the scan mode (m/z 50-650) and the same sample analyzed with SIM mode. A. Total ion chromatogram (TIC), B. the common base ion m/z 316 for all AR homologues, C. molecular ions m/z 540 for C17:0, D. molecular ions m/z 568 for C19:0, F. molecular ions m/z 582 for C20:0, F. molecular ions m/z 596 for C21:0, G. molecular ions m/z 624 for C23:0, and H. molecular ions m/z 652 for C25:0.
Table 10. Intra- and inter-batch precision (as coefficient of variation) for determination of alkylresorcinols (AR) in adipose tissue and a plasma sample used as in-house quality controls (n=4 per batch, 3 batches in total) and recovery from spiking experiments (n=4 for each spiking level). (CV=coefficient of variation)

<table>
<thead>
<tr>
<th>AR in</th>
<th>Average</th>
<th>Intra-batch CV (%)</th>
<th>Inter-batch CV (%)</th>
<th>Recovery, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Intra-batch</td>
<td>Inter-batch</td>
<td>Low¹</td>
</tr>
<tr>
<td>Adipose</td>
<td></td>
<td>CV (%)</td>
<td>CV (%)</td>
<td></td>
</tr>
<tr>
<td>C17:0</td>
<td>5</td>
<td>13.7</td>
<td>16.4</td>
<td>109</td>
</tr>
<tr>
<td>C19:0</td>
<td>31</td>
<td>4.4</td>
<td>8.3</td>
<td>88</td>
</tr>
<tr>
<td>C21:0</td>
<td>87</td>
<td>3.5</td>
<td>11.3</td>
<td>61</td>
</tr>
<tr>
<td>C23:0</td>
<td>53</td>
<td>4.4</td>
<td>14.2</td>
<td>81</td>
</tr>
<tr>
<td>C25:0</td>
<td>89</td>
<td>4.1</td>
<td>12.0</td>
<td>105</td>
</tr>
<tr>
<td>Plasma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C17:0</td>
<td>5.7</td>
<td>3.3</td>
<td>5.9</td>
<td>110</td>
</tr>
<tr>
<td>C19:0</td>
<td>17.2</td>
<td>1.8</td>
<td>8.5</td>
<td>97</td>
</tr>
<tr>
<td>C21:0</td>
<td>18.8</td>
<td>2.6</td>
<td>12.4</td>
<td>90</td>
</tr>
<tr>
<td>C23:0</td>
<td>6.6</td>
<td>3.2</td>
<td>2.6</td>
<td>83</td>
</tr>
<tr>
<td>C25:0</td>
<td>7.0</td>
<td>7.5</td>
<td>8.8</td>
<td>99</td>
</tr>
</tbody>
</table>

¹ Spiked amount of each AR homologs was 2 ng for plasma samples and 0.4 ng for adipose tissue samples.
² Spiked amount of each AR homologs was 4 ng for plasma samples and 2 ng for adipose tissue samples.
³ Spiked amount of each AR homologs was 6 ng for plasma samples and 6 ng for adipose tissue samples.
4.2 Alkylresorcinols in adipose tissue in response to a 12-week controlled whole grain intervention

In Paper II, the response of AR concentrations in adipose tissue to controlled whole grain intake under intervention conditions for 12 weeks was evaluated for the first time. At baseline (week 0), there was no difference in whole grain intake between the groups (Figure 7). The estimated whole grain intake at baseline was in agreement with intakes for the general Finnish population reported in the National FINDiet 2012 survey (Helldán et al., 2013).

During the 12 weeks intervention, WGDG group did not change its whole grain intake significantly compared with baseline, due to already high habitual whole grain intake at baseline, but RDG significantly reduced its whole grain intake during the RDG intervention.

The AR concentrations in plasma and adipose tissue were in the range reported previously for Nordic populations (Aubertin-Leheudre et al., 2008, Jansson et al., 2010, Kyrö et al., 2014a, Landberg et al., 2011, Magnusdottir et al., 2013). In accordance with the reported whole grain intake, AR in plasma or adipose tissue at baseline did not differ between the groups.

![Figure 7. Median and 25th-75th percentiles (error bars) of whole grain intake reported in the whole grain diet group (WGDG) and refined diet group (RDG) at week 0 and week 12. Wilcoxon signed-rank test was used to compare each parameter within group between week 0 and week 12, and Mann-Whitney’s U-test for parameters between treatment groups. *p<0.05, **p<0.01 and ***p<0.001.](image-url)
Figure 8. Median and 25th-75th percentiles (error bars) of alkylresorcinol (AR) concentrations in plasma and adipose tissue at week 0 and week 12 among subjects in the whole grain diet group (WGDG) and refined diet group (RDG). Wilcoxon signed-rank test was used to compare each parameter within group between week 0 and week 12. *p<0.05, **p<0.01 and ***p<0.001.
After the intervention, concentrations of all AR homologs in adipose tissue and plasma differed between treatments (Figure 8). This suggest that both plasma and adipose tissue reflected the reduction in whole grain intake in RDG during the 12-week intervention.

Average whole grain intake from three 4-day food records (week 3, week 7 and week 11) was strongly correlated with AR in adipose tissue \( (r=0.60-0.84, P<0.05) \) and in plasma \( (r=0.60-0.72, P<0.05) \) at week 12. However, there was no significant association between whole grain intake and AR in adipose tissue or in plasma at baseline, which might due to lack of precision in estimating whole grain intake at base-line, where subjects had their habitual diet.

4.3 Alkylresorcinols in adipose tissue as biomarkers under free-living conditions

4.3.1 Relative validity and determinants beyond intake

*Paper III*

In *Paper III*, the performance of AR in adipose tissue as biomarkers of long-term whole grain intake was investigated among free-living Swedish men and women. At the time of adipose and plasma sampling, men were older than women, but BMI was similar between men and women. Adipose tissue from men contained a higher concentration of total AR than adipose tissue from women, while AR C17:0/C21:0 ratio in adipose tissue, total AR concentration, and AR C17:0/C21:0 ratio in plasma did not differ significantly between men and women.

The concentration of AR in adipose tissue was within the range reported previously for a few SMC participants (Jansson *et al.*, 2010) and for Finnish men and women (Paper II).

Similar whole grain intake was reported in the three repeated FFQs among men over a period of 17 years (Figure 9). For women, significantly higher whole grain intakes \( (P<0.05) \) and lower energy intake \( (P<0.05) \) were reported in FFQ-87 compared with the other FFQs, but no difference was found between FFQ-97 and FFQ-03.
Figure 9. Median, 25th percentile, and 75th percentile of daily whole grain wheat (WGW) and whole grain rye (WGR) intake from food frequency questionnaires (FFQ) for women (in 1987, 1997 and 2003) and men (in 1997, 2009 and 2010). Intra-class correlations (ICC) of WGW and WGR between the closest FFQ are also listed. *, intake variable significantly higher than the corresponding intake from other FFQ in the Student’s t-test for the same gender.

Figure 10. Percentage change in total alkylresorcinol (AR) concentration in adipose tissue and plasma per 10 g/day increment in energy-adjusted (residual method) intake of whole-grain wheat (WGW) and whole-grain rye (WGR), per 1 Mcal/day increment in total energy intake, sex (female vs male), per 1 kg/m² increment in body mass index (BMI), and per year increment in age among free living Swedish men (n=149) and women (n=109). Plasma samples were missing from 19 women. Model 1 included intake of WGW and WGR, and total energy intake; Model 2 included intake of WGW and WGR, total energy intake, sex, body mass index (BMI), and age.
The correlations between whole grain intakes and AR in adipose tissue were modest, with TWG ($r=0.22-0.42$) and WGR&W ($r=0.27-0.44$) were more strongly correlated to biomarkers than whole grain products intake ($r=0.15-0.36$), illustrating the advantages of deriving and using whole grain intakes (g/d) instead of product intake estimates in this population.

Correlation coefficients were strongest for WGR ($r=0.31-0.42$), weaker for WGR&W and TWG, and poor for WGW ($r=0.09-0.33$). This is most likely because WGR is mainly from crispbread which is consumed frequently and regularly (contributing 67-74% of WGR; ICC: 0.37-0.56); whereas the main sources of WGW are whole-meal bread and pasta (contributing 58-73% of WGW), which have larger variation in whole grain contents and are more irregularly consumed (ICC: 0.01-0.28). Correlation coefficients for TWG reported and AR concentration in adipose tissue in this study were in agreement with those found previously in one small study in the same population where an FFQ was used (Jansson et al. 2010), but much weaker than when using whole grain intake estimated by food records ($r=0.60-0.84$) (Paper II). This suggests that FFQ-based whole grain intake values are more prone to measurement errors than values from food records.

Repeated FFQs could be used to improve precision in estimating long-term whole grain intake and thereby improve correlations, as shown in studies comparing estimated fatty acid intakes by repeated FFQs with corresponding fatty acids in erythrocytes (Sun et al., 2007). In Paper III, correlation coefficients between WGW or WGR and adipose tissue AR concentrations were stronger when the mean of repeated FFQs was used rather than estimates from a single FFQ, even when that FFQ was close in time to the biopsy. Correlations between WGR from repeated FFQs and biomarkers were improved to a lesser extent than for WGW, probably due to better precision in estimating long-term WGR intake.

Correlations between whole grain intake and biomarkers were stronger for women than men, suggesting that women had more stable whole grain intake than men and/or reported their intakes more accurately, despite longer time intervals between their repeated FFQs.

Among non-dietary determinants investigated, gender was found to be significantly associated with total AR concentration in adipose tissue (Figure 10). Adipose tissue total AR concentration was 61.3 % lower in women than in men after adjusting for energy-adjusted whole grain wheat and rye intake, BMI, and age.

Similarly, higher plasma AR concentrations have been found for men than women in several previous studies (Montonen et al., 2010, Ross et al., 2012) and the difference has been suggested to be due to faster elimination of AR in women than in men (Frank et al., 2008). Higher plasma AR concentrations among men than women are likely to lead to more extensive accumulation in adipose tissue in men. An
additional reason for lower adipose tissue AR concentration in women could be larger distribution volume in women due to higher body fat content among women than men (Ley et al., 1992). However, this cannot be confirmed since X-ray absorptiometry data were not available for the men studied in Paper III. In contrast to what has been shown for AR in plasma in other populations, there was no evidence of a difference in AR concentrations in adipose tissue (or in plasma) due to age.

**Paper V**

In Paper V, total AR concentration in adipose tissue of the Danish subjects was within the same range as observed among Swedish subjects in Paper III and as reported in a previous study (Jansson et al., 2010), but lower than in Finnish subjects (Paper II). This is in agreement with the whole grain intake reported in these three countries (Finland>Sweden=Denmark) (Kyrø et al., 2012, Nurmi et al., 2009). The AR C17:0/C21:0 ratio in adipose tissue from the Danish women in Paper V was lower than among Swedish subjects (Paper III). The slightly lower AR C17:0/C21:0 ratio in adipose tissue is in contrast to the higher proportion of whole grain rye in the diet observed among the Danish women compared with Swedish participants in Paper III. In Paper V, total AR concentration in adipose tissue was modestly correlated with WGR&W ($r=0.31, P<0.001$), WGR ($r=0.28, P<0.001$) and WGW intake ($r=0.16, P<0.001$). The correlation coefficients were similar to those found for Swedish subjects (Paper III).
4.4 Alkylresorcinols in adipose tissue vs. alkylresorcinols in plasma as biomarkers

Based on the results in Paper II, it appears that AR in adipose tissue have a much faster turnover rate than many other lipophilic compounds, such as fatty acids (Beynen et al., 1980, Handelman et al., 1994, Strawford et al., 2004). Lipophilic compounds are typically stored in adipocyte cytoplasmic lipid droplets (Brasaemle, 2007) but, in analogy with tocopherols, AR are known to associate in biological membranes (Asghar et al., 1991, Linko and Adlercreutz, 2005, Simon et al., 2001). We hypothesized that AR are incorporated into the membranes of adipocytes (Asghar et al., 1991, Linko and Adlercreutz, 2005, Simon et al., 2001) where the turnover rate is faster than for lipophilic compounds in adipocyte cytoplasmic lipid droplets (Beynen et al., 1980, Handelman et al., 1994, Strawford et al., 2004). The suggested rapid turnover rate of AR in adipose tissue may explain why they are well correlated with AR in plasma (r=0.40-0.71, P<0.05 for AR C17:0, C19:0, C23:0 and C25:0; r=0.27, P=0.17 for AR C21:0).

Initially, due to longer anticipated half-life of AR in adipose tissue than in plasma (Strawford et al., 2004), AR concentration in adipose tissue were expected to be better correlated to FFQ-based whole grain intakes than AR in plasma in Paper III. However, the correlations were of similar magnitude. One reason could be that AR in adipose tissue is located in the adipocyte membrane and that the turnover is rapid, as discussed above. Another reason could be that plasma AR concentration also reflected long-term whole grain intakes among the subjects in Paper II and III, as has been shown from populations where whole grain intake is frequent and regular (Andersson et al., 2011, Landberg et al., 2013, Montonen et al., 2010). Moreover, the large measurement error inherent in FFQ-based whole grain intakes might have masked any potential difference in correlations with AR in adipose tissue and plasma in Paper III. The fact that the strongest correlation between AR in adipose tissue and plasma was observed between short homologs and that AR concentrations in adipose tissue were more than 10-fold higher than in fasting plasma suggests that the AR concentration in fasting plasma may to a large extent derive from AR liberated back from adipose tissue, as has been shown for fatty acids in fasting plasma (Halliwell et al., 1996). Mechanistic, controlled feeding studies needs to be conducted to clarify what mechanism that is at play.

Based on the results from Papers II-IV, it appears that AR concentrations in adipose tissue reflect long-term whole grain rye and wheat intake in a similar way as AR concentrations in plasma, in the populations investigated in this thesis. A clear advantage of using AR in adipose tissue biopsies instead of AR in plasma is that the former are not likely to be affected and potentially confounded by the lipoproteins that transport AR in plasma. On the other hand, adipose biopsy samples are less available from studies compared with plasma samples.
4.5 Alkylresorcinols in adipose tissue in relation to blood lipids

The associations between whole grain wheat and rye intake estimated by FFQ or reflected by AR in adipose tissue or plasma with blood lipid concentrations were investigated using data and samples from women in Paper III. BMI and age did not significantly differ across tertiles of AR (T1-T3) in adipose tissue and in plasma. As mentioned earlier, AR concentrations in adipose tissue and plasma and the AR C17:0/C21:0 ratio in Paper III were similar to those reported in other Nordic studies with mixed consumption of whole grain wheat and rye (Andersson et al., 2011, Cuff et al., 2015, Jansson et al., 2010, Kyro et al., 2014a, Landberg et al., 2009c, Marklund et al., 2013, Nybacka et al., 2016).

Total AR concentration in adipose tissue was inversely associated with TCHOL and HDL-C (P-trend<0.05) (Figure 11). Participants in T3 of total AR concentration in adipose tissue had 8.7% (95% CI: 1.3-15.4%) lower TCHOL and 13.5% (95% CI: 2.9-23.0%) lower HDL-C compared with participants in T1. These findings suggest that total intake of whole grain rye and wheat was associated with lower TCHOL, mainly driven by lower HDL-C concentrations.

Separate roles of whole grain rye and whole grain wheat were investigated using AR C17:0/C21:0 ratio in adipose tissue (Figure 11). A higher ratio of AR C17:0/C21:0, reflecting higher relative whole grain rye intake, was associated with higher HDL-C and lower LDL-C (P-trend<0.05). Participants in T3 of the AR C17:0/C21:0 ratio in adipose tissue had a 14.2% (95% CI: 1.6-28.4%) higher HDL-C and a 13.9% (95% CI: 3.0-23.5%) lower LDL-C compared with participants in T1.

Based on the results for total AR concentration and AR C17:0/C21:0 ratio in adipose tissue, it appears that high total and relative whole grain wheat intake may explain the inverse association with HDL-C observed. The lower LDL-C associated with higher C17:0/C21:0 ratio is in agreement with a previous study, where this ratio in plasma was associated with lower LDL-C in men and women with signs of metabolic syndrome (Magnusdottir et al., 2014c).

We hypothesize that the beneficial association observed for whole grain rye, reflected by the C17:0/C21:0 ratio, is due to higher content of viscous dietary fiber (Buksa et al., 2016, Laerke et al., 2015, Pettersson and Åman, 1988) and phytosterols (Andersson et al., 2013, Nurmi et al., 2008) in rye than in wheat. Around 80% of the viscosity in rye has been attributed to water soluble arabinoxylans (Bengtsson et al., 1992). Although arabinoxylans can be found in both whole grain rye and whole grain wheat, the viscosity formed by
Figure 11. Percentage change and P-trend (*P<0.05, **P<0.01 and ***P<0.001) of total cholesterol, high density lipoprotein-cholesterol (HDL-C), low density lipoprotein-cholesterol (LDL-C), and triglycerides (TG) across tertiles of total whole grain rye and wheat intake or relative whole grain rye intake derived from food frequency questionnaire (FFQ), alkylresorcinols (AR) in adipose tissue, or in plasma among free living Swedish women (n=109). Plasma samples were missing from 19 women. The multivariate models were adjusted for body mass index (BMI) and age. WGR&W = sum energy-adjusted intake of whole grain rye and whole grain wheat from FFQ; WGR% = the ratio of FFQ-based whole grain rye intake and WGR&W.
Arabinoxylans in rye is more extensive and resistant to food processing (Johansson et al., 2017). Pigs fed rye have significantly higher viscosity of ileal digesta than pigs fed whole grain wheat (Laerke et al., 2015). The resistant viscosity formed by arabinoxylans from rye might lower LDL-C in the blood through similar mechanisms as beta-glucan from oats and barley (Wolever et al., 2010). Future dietary interventions with different doses of rye foods rich in viscosity forming arabinoxylans should be conducted to test this hypothesis.

4.6 Alkylresorcinols in adipose tissue in relation to breast cancer incidence

In Paper V, there was no statistically significant association between WGR&W intake or AR concentration in adipose tissue and the risk of developing breast cancer (total, ER+, or ER-) (Figure 12). This is in agreement with findings from a prospective analysis based on the entire “Diet, Cancer and Health” cohort (Egeberg et al., 2009) and suggests that total whole grain rye and wheat intake has a limited preventive effect on associated breast cancer incidence among women in a population with high and diverse whole grain intake. This is supported by results from observational studies (Suzuki et al., 2008, Terry et al., 2002, Zhang et al., 2011) which in general show no significant associations between cereal fiber intake and breast cancer. However, some bioactive compounds found in whole grain rye and wheat have been suggested to reduce breast cancer development (Adlercreutz, 2010, Jacobs et al., 1995, Slavin, 2000).

One of the most studied group of bioactive compounds in this aspect are the lignans (Adlercreutz et al., 1992, Adlercreutz et al., 1993, Adlercreutz, 1998, Adlercreutz, 2007, Andersson et al., 2013, Nilsson et al., 1997, Slavin, 2003). Plant lignans are metabolized by intestinal bacteria to the mammalian lignans enterodiol and enterolactone (Jacobs et al., 2007). Enterolactone has been suggested to reduce the risk of breast cancer by lowering the activity of estrogen (Adlercreutz et al., 1992, Adlercreutz et al., 1993, Adlercreutz, 1998, Adlercreutz, 2007), and has been found to be inversely associated with ER+ breast cancer, but to have no effect on ER- breast cancer, in tamoxifen prevention trials (Cuzick et al., 2003).

In a previous study with all Danish postmenopausal women from the same cohort as used in Paper V, an inverse association between circulating enterolactone concentrations and risk of developing breast cancer was found for ER-, but not ER+, breast cancer (Olsen et al., 2004). However, enterolactone is affected by several determinants other than whole grain wheat and rye, which may have contributed to the enterolactone concentrations associated with reduced ER- breast cancer risk in that study.
Another reason for the lack of association between total whole grain wheat and rye intake and breast cancer could be differing roles of rye and wheat in modifying the risk of developing breast cancer.

![Figure 12: Association between total whole grain rye and wheat intake or relative whole grain rye intake and risk of developing invasive breast cancer (BC), estrogen-receptor-positive (ER+), or estrogen-receptor-negative (ER-). AR, alkylresorcinols; WGR&W= sum of energy adjusted intake of whole grain rye and whole grain wheat from FFQ; WGR%=ratio of FFQ-based whole grain rye intake to WGR&W. Hazard ratios (HR) and 95% Confidence intervals of breast cancer for quartiles of total whole grain rye and wheat intake (WGR&W or total AR concentration in adipose tissue), or relative whole grain rye intake (WGR% or AR C17:0/C21:0 ratio). HR for per unit of these exposures (10g/day for WGR&W; 10% for WGR%; 1nmol/L for total AR concentration; and 0.1 for AR C17:0/C21:0 ratio) were also calculated.]

In **Paper V**, a higher AR C17:0/C21:0 ratio in adipose tissue, which partly reflected a higher relative whole grain rye (or lower relative whole grain wheat) intake was consistently associated with an increased risk of ER+, but not ER-breast cancer (Figure 12). Although *P*-trends for WGR% did not reach statistical significance, women in the highest quartile of WGR% had around 55% higher risk of ER+ breast cancer than women in the lowest quartile. Higher AR 17:0/C21:0 ratio in adipose tissue could reflect higher whole grain rye intake in the diet, but might also relate to as yet unknown dietary/non-dietary factors.
A high intake of short AR homologs such as AR C17:0, which is mainly from whole grain rye (Chen et al., 2004, Menzel et al., 2012, Ross et al., 2003c) will result in a high C17:0/C21:0 ratio. Theoretically, the presence of C17:0 in adipose tissue of the breast may have promoted synthesis of progestogens and suppressed the downstream metabolism of progestogens and estradiol, as shown in vitro (Oskarsson and Ohlsson Andersson, 2016). High concentrations of these sex hormones are strongly associated with elevated risk of postmenopausal breast cancer (Key et al., 2002). In support of this theory, it was indicated in Paper V that high concentrations of AR C17:0, but not AR C19:0 or ARC21:0 were associated with an increased risk of developing ER+ breast cancer (data not shown). Future intervention studies in animals and humans are needed to evaluate the potential role of AR on sex-hormone metabolism and the role of C17:0/C21:0 as a biomarker in relation to breast cancer.
5 General discussion

Alkylresorcinols concentrations in plasma and the concentration of their metabolites in plasma and urine have been evaluated and increasingly used as dietary biomarkers in epidemiological and whole grain intervention studies (Biskup et al., 2016, Drake et al., 2013, Landberg et al., 2008a, Magnusdottir et al., 2013, Magnusdottir et al., 2014a, Magnusdottir et al., 2014b, McKeown et al., 2016b, Nybacka et al., 2016, Olsen et al., 2010, Ross et al., 2004d). However, due to the short half-life, their use may be limited to populations where the intake of whole grain is high and stable (Andersson et al., 2011, Landberg et al., 2013, McKeown et al., 2016a, Montonen et al., 2010). This PhD project was based on the hypothesis that this could be overcome by evaluating AR concentration in adipose tissue as more long-term biomarkers. AR were found to be present in adipose tissue of Swedish women in a small pilot project conducted before the start of this project and the concentrations were found to be correlated with estimated intake (Jansson et al., 2010). However, the responsiveness of AR in adipose tissue to controlled intake, factors affecting adipose tissue AR concentrations, and the performance of AR in adipose tissues as biomarkers and their relationship with disease endpoints had never been tested.

It took around two years and more than 10,000 sample injections to develop and evaluate the analytical method for analysis of AR in adipose tissue (Paper I). The GC-MS method developed can be used for both adipose tissue and plasma samples. Moreover the sensitivity is 10-to 40-fold higher and the method is 3 ×faster in running time, compared with the original method suggested for analysis of AR in plasma and adipose tissue (Jansson et al., 2010, Landberg et al., 2009b). Around 100 adipose tissue or plasma samples per day can be analyzed with the new method, which enabled large-scale validations and applications in the studies on which this thesis is based (Paper III-V) and in studies to come.
Since a ‘gold standard’ for determination of whole grain intake is still lacking (Kaaks, 1997), in this thesis evaluation of AR concentrations in adipose tissue as biomarker of whole grain wheat and rye intake was carried out by assessing the biomarkers with whole grain intake precisely measured by repeated food records during an intervention over a relatively short period of time (12 weeks) (Paper II), and with long-term intake of whole grain rye and wheat from single (Paper III-V) and repeated FFQs (Paper III). The latter two methods are less precise (Kristal et al., 2005), but more practical and frequently applied in observational studies (Johansson, 2014). The AR concentration in adipose tissue were found to be modestly correlated with total whole grain wheat and rye intake or whole grain rye intake, but weakly correlated with whole grain wheat intake, most probably due to poor estimation of whole grain wheat in unspecific FFQs. These findings suggest that there is room for improvement on whole grain wheat intake estimation with FFQ, and specific FFQs with more cereal food items could be a good option (Ross et al., 2009). It was found that sex affected AR concentrations in adipose tissue (Paper III) and therefore any analysis using AR in adipose tissue as biomarkers should be adjusted for sex.

In comparison with plasma AR measured in populations with high and frequent whole grain intake (Paper II-V), it appears that AR in adipose tissue in AR reflect whole grain wheat and rye in a similar way. This may be due to a rapid exchange between AR in plasma and adipose tissue and/or due to the accurate reflection of long-term AR concentrations in plasma in these populations.

Since adipose tissue AR concentrations appeared to reflect whole grain wheat and rye in a corresponding way to plasma AR concentrations, their associations with relevant endpoints were investigated for the first time (Papers IV and V).

A consistent inverse association between AR C17:0/C21:0 ratio in adipose tissue or plasma and LDL-C was found in Paper IV which is in agreement with results from another study and suggests that whole grain rye rich in viscous fiber may have a role in cholesterol reduction. This should be confirmed in dietary intervention studies and may open new paths for development of healthy food products with health claims.

The role of whole grain and cereal fiber in breast cancer prevention is controversial. It has been hypothesized, but not proven, that rye, which is also rich in lignans, may be of particular benefit. This thesis is the first study to investigate the association, using objective biomarkers of whole grain intake. It was surprising to find a strong and positive association between higher AR C17:0/C21:0 ratio in adipose tissue and risk of developing ER+ breast cancer (Paper V). Further studies are needed to replicate this finding in independent cohorts and to understand the underlying mechanisms.
The two endpoint studies in this thesis support the idea that whole grain wheat and rye may have partly different effects on health and that future nutritional studies should put more emphasis on the role of separate grains on human health. Dietary biomarkers could provide essential and complementary information in such analyses and should be used in parallel with conventional dietary assessment methods.
6 Main findings

1. A high throughput GC-MS method for determination of alkylresorcinol concentrations in adipose tissue biopsies (10-50 mg) was developed, evaluated (Paper I), and applied in different studies (Papers II-V).

2. Total alkylresorcinol concentration in adipose tissue responded to reduced whole grain intake over a 12wk intervention and correlated well with whole grain intake estimated by 3×4-day food records (Paper II).

3. Total alkylresorcinol concentration in adipose tissue was correlated with long-term whole grain intake assessed over 14-17 years in free-living Swedish men and women. Sex was found to be an important determinant of total AR concentration in adipose tissue (Paper III).

4. Alkylresorcinol concentrations in adipose tissue correlated well with fasting plasma alkylresorcinol concentrations in the investigated population studied, which have high and frequent whole grain intake (Paper II-III).

5. Higher AR C17:0/C21:0 ratios in adipose tissue and plasma, reflecting higher proportions of whole grain rye in total whole grain wheat and rye intake, were consistently associated with lower low-density lipoprotein cholesterol (LDL-C) in blood. These findings are supported by plausible mechanisms of LDL-C reduction, but should be confirmed in randomized controlled trials (Paper IV).

6. Total whole grain wheat and rye did not appear to be related to risk of developing breast cancer. The AR C17:0/C21:0 ratio in adipose tissue was positively associated with risk of ER+ breast cancer, but the exact mechanism(s) behind this observation remains to be elucidated (Paper V).
7 Future research

1. Total alkylresorcinol concentration in adipose tissue should be further evaluated as long-term biomarkers of whole grain wheat and rye intake in subjects with low and infrequent whole grain intake. In such studies, whole grain intake should be estimated with methods with higher precision, such as repeated weighed food records or specific FFQ with more cereal items.

2. Effects of age and body fat composition on alkylresorcinol concentrations should be further evaluated in populations with broader age range and body fat measurements with dual energy X-ray absorptiometry for both genders.

3. The elimination half-life of different alkylresorcinol homologs in adipose tissue should be estimated, preferably in populations with no/minimal background exposure to alkylresorcinols.

4. Determinants of AR C17:0/C21:0 ratio in adipose tissue should be established under controlled intervention conditions.

5. Effects of whole grain rye intake on blood lipid profile should be confirmed in human intervention studies.

6. Findings in Paper V should be replicated in an independent population and the effects of whole grain rye intake on sex hormone metabolism should be examined in animal or human intervention studies.
References


WHO 2013. Follow-up to the political declaration of the high-level meeting of the General Assembly on the Prevention and Control of Non-Communicable Diseases. WHA6610: WHO.


Acknowledgements

This work was mainly carried out at the former Department of Food Science, SLU, Uppsala, Sweden, with samples and support proved by different Nordic groups from Karolinska Institute, Chalmers University of Technology, University of Eastern Finland, and Danish Cancer Society. I would like to thank all the men and women who participated in these studies and also to express my thanks for financial support from FORMAS, the Swedish Research Council-Medicine, and the Swedish Nutrition Foundation.

Rikard Landberg (main supervisor) – thank you for your support during these five years of my PhD studies. This work would not have started without your insistence on recruiting me to your group, and would not have finished so well without your dedicated help. I really enjoyed the freedom and understanding you provided during this journey of research. Alicja Wolk (co-supervisor) – thank you for conducting the excellent SMC&COSM studies and selflessly sharing your ideas, precious samples and valuable data, which have been collected since the year I was born. Ali Moazzami (co-supervisor) – although the work I am doing is a little far away from your area, I am grateful for your suggestions from time to time. It was so much fun to discuss with you that we almost missed the plane at Aarhus airport. Janicka Nilsson – without your support during these five years, the projects would not have run so smoothly, especially after the move to Chalmers. You showed me how a good lab is organized. It is more fun to work in a lab with everything ready to go. Roger Andersson – thank you for discussions on all kinds of topics, from dietary fiber in cereals, GC-FID, computers in the old days, to SAAB cars. You are always helpful and kind to everyone. Galia Zamaratskaia – I really enjoyed the food you shared, and the nice time we had with all the other LiFT members in Beijing. Carl Brunius – thank you for all the discussions about LC-MS. Anne Kirstine, Kia, Lin, Adila, Roksana, Lillie, and Izabela, it has been so nice to work with you in the Food and Health group. Many thanks also to Otto Savolainen for all the help during
method development at Chalmers. I would like to thank Anja Olsen and Cecilie Kyro for hosting and advising me during the data processing at the Danish Cancer Society Research Center. Jana Pickova – thank you for sharing your knowledge about GC-FID and fatty acid synthesis. Åse Lundh – thank you for introducing me to the interesting world of milk. Many thanks to Kristine Koch, Lena Dimberg, Monika Johansson, Annica Andersson, Gunnul Fransson, Camilla Löfgren, Jaana Evander, Maggan, and Carina for all the help during the years. Kerstin Lundström & Anders H Karlsson – thank you for introducing me to the wonderful world of meat science, which inspired me greatly during the development of sample preparation for Paper I.

To my PhD mates: Daniel Johansson, Susanne Djurle, Zhao Xue, Maria Karlsson, Isabella Lee, and Anna Sandberg, thank you for the nice and warm office environment. Sun Li & Huang Siming, Wang Guodong, & Li Yajun, Liu Tong & Zeng Xin, Lu Jing & Wang Pengcheng, Cheng Ken & Sun Xu, Liu Shujing & Huang Zhiqiang, Jin Yunkai & Liu Qinsong, Jiang Hua & Yi Jun, Zhao Ming, Liang Qin, Feng Xinmei, Hu Guozhen and all other friends, thank you for all the parties and hot pots.

致 吴和平 与 李翠瑛：从牙牙学语到如今独当一面，感谢近三十年来的养育与支持。感谢 董天怀 与 刘莉 多年来对女婿的照顾与包容。最后感谢我的妻子 董文博， 从签证相遇到结婚，一路走来相知相爱，愿得一人心，白首不相离。