

(19)



SUOMI - FINLAND
(FI)

PATENTTI- JA REKISTERIHALLITUS
PATENT- OCH REGISTERSTYRELSEN
FINNISH PATENT AND REGISTRATION OFFICE

(10) **FI 130157 B**

(12) **PATENTTIJULKAISU**
PATENTSKRIFT
PATENT SPECIFICATION

(45) Patentti myönnetty - Patent beviljats - Patent granted **21.03.2023**

(51) Kansainvälinen patenttiluokitus - Internationell patentklassificering - International patent classification
G01N 33/68 (2006 . 01)
G01N 33/92 (2006 . 01)
G01N 24/08 (2006 . 01)

(21) Patenttihakemus - Patentansökan - Patent application **20215110**

(22) Tekemispäivä - Ingivningsdag - Filing date **01.02.2021**

(23) Saapumispäivä - Ankomstdag - Reception date **01.02.2021**

(41) Tullut julkiseksi - Blivit offentlig - Available to the public **02.08.2022**

(73) Haltija - Innehavare - Holder

1• Nightingale Health Oyj, c/o Nightingale Health Oy Mannerheimintie 164a, 00300 HELSINKI, (FI)

(72) Keksijä - Uppfinnare - Inventor

1• Würtz, Peter, HELSINKI, (FI)
2• Julkunen, Heli, HELSINKI, (FI)

(74) Asiamies - Ombud - Agent

Papula Oy, P.O. Box 981, 00101 Helsinki

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention

Menetelmä rintasyöpään kuolemisen riskin määrittämiseksi kohteelle
Förfarande för att bestämma om en person har risk att dö i bröstcancer
METHOD FOR DETERMINING WHETHER A SUBJECT IS AT RISK OF DYING FROM BREAST CANCER

(56) Viitejulkaisut - Anförda publikationer - References cited

JOBARD, E. ET AL. A serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer, 2014-02-01, Cancer Letters, Vol. 343, No. 1, pages 33-41; SUMAN, S. ET AL. Metabolic fingerprinting in breast cancer stages through 1 H NMR spectroscopy-based metabolomic analysis of plasma, 2018-10-25, Journal of Pharmaceutical and Biomedical Analysis, Vol. 160, pages 38-45; US 2018112274 A1; DUPREZ, D. A. ET AL. Comparison of the Predictive Value of GlycA and Other Biomarkers of Inflammation for Total Death, Incident Cardiovascular Events, Noncardiovascular and Noncancer Inflammatory-Related Events, and Total Cancer Events, 2016-07-01, Clinical Chemistry, Vol. 62, No. 7, pages 1020-1031; GRUPPEN, E. G. ET AL. GlycA, a novel pro-inflammatory glycoprotein biomarker is associated with mortality: results from the PREVEND study and meta-analysis, 2019-11-30, Journal of Internal Medicine, Vol. 286, No. 5, pages 596-609; WO 2013184483 A1; THOMPSON D.K. ET AL. Elevated serum acute phase protein levels as predictors of disseminated breast cancer, 1983-06-01, Cancer, Vol. 51, No. 11, pages 2100-2104; WO 2011157446 A1; EP 2481816 A1; WO 2019233028 A1

(57) Tiivistelmä - Sammandrag - Abstract

Esitetään menetelmä sen määrittämiseksi, onko kohteella riski kuolla rintasyöpään tai eturauhassyöpään.

A method for determining whether a subject is at risk of dying from breast cancer or prostate cancer is disclosed.

C50: Malignant neoplasm of breast

Mortality

n = 63 859
311 events

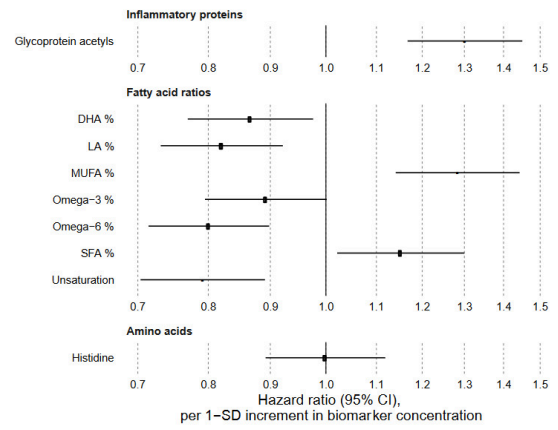


Figure 1

METHOD FOR DETERMINING WHETHER A SUBJECT IS AT RISK OF DYING FROM BREAST CANCER

TECHNICAL FIELD

5 The present disclosure relates generally to methods for determining whether a subject is at risk of dying from breast cancer.

BACKGROUND

10 Breast and prostate cancers are common cancers and frequent causes of cancer related deaths among women and men, respectively. They also place a great strain on the health care system and have a significant social and financial impact on the affected patients and their
15 families. Fortunately, due to the improved treatment options, these cancers are considered potentially curable when detected at an early stage and it is thus possible to influence the disease burden.

 Prediction of breast cancer and prostate cancer
20 survival and an accurate identification of individuals at high risk of dying from these cancers later in life may be important to support early treatment and prevent the development of more severe presentations of these diseases. Various biomarkers may be useful for
25 predicting breast and prostate cancer survival and/or whether an individual person is at an elevated risk of dying from these cancers in the future. Such biomarkers may be measured from various biological samples, for example from blood samples or related biological fluids.
30 Biomarkers predictive of the severe presentations of and risk of mortality from these cancers could support targeted screening, earlier diagnosis and treatment, including guidance for deciding on prophylactic mastectomy. Such a method could have utility both for
35 screening general population without known diagnosis of breast or prostate cancer, as well as for predicting the prognosis and survival among individuals diagnosed with

breast cancer or prostate cancer to guide treatment options.

Jobard et al., *Cancer Letters* 2014, 343, 1, 33-41 describes a serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer.

Suman et al., *Journal of Pharmaceutical and Biomedical Analysis* 2018, 160, 38-45 describes metabolic fingerprinting in breast cancer stages.

10 US 2018/112274 discloses methods and systems for predicting colorectal cancer incidence and mortality.

Duprez et al., *Clinical Chemistry* 2016, 62, 7, 1020-1031 describes a comparison of the predictive value of GlycA and other biomarkers of inflammation for total death, incident cardiovascular events, noncardiovascular and noncancer inflammatory-related events, and total cancer events.

Gruppen et al., *Journal of Internal Medicine* 2019, 286, 5, 596-609 describes the determination of the association of GlycA with overall and cause-specific mortality in a cohort of men and women.

WO 2013/184483 discloses NMR measurements of GlycA.

25 Thompson et al., *Cancer* 1983, 51, 11, 2100-2104 describes elevated serum acute phase protein levels as predictors of disseminated breast cancer.

WO 2011/157446 describes biomarkers for the prediction of incident cancer.

30 EP 2481816 describes markers to predict survival of breast cancer patients and uses thereof.

WO 2019/233028 discloses a biomarker and a method for diagnosing or predicting the risk of death.

SUMMARY

A method for determining whether a subject is at risk of dying from breast cancer is disclosed. The method may comprise determining in a biological sample obtained from the subject a quantitative value of at least one biomarker of the following:

- glycoprotein acetyls,
- a ratio of docosahexaenoic acid to total fatty acids,
- a ratio of linoleic acid to total fatty acids,
- a ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids,
- a ratio of omega-3 fatty acids to total fatty acids,
- a ratio of omega-6 fatty acids to total fatty acids,
- a ratio of saturated fatty acids to total fatty acids,
- fatty acid degree of unsaturation,
- histidine; and

comparing the quantitative value(s) of the at least one biomarker to a control sample or to a control value;

wherein an increase or a decrease in the quantitative value(s) of at least one biomarker, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of dying from breast cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are included to provide a further understanding of the embodiments and constitute a part of this specification, illustrate various embodiments. In the drawings:

Figure 1 shows the relation of baseline concentrations of 9 blood biomarkers to breast cancer mortality (malignant neoplasm of breast, ICD-10 diagnosis code C50 in death records) in a general population screening setting (such as along with a mammography test), when the biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population.

Figure 2 shows the relation of baseline concentrations of 9 blood biomarkers to prostate cancer mortality (malignant neoplasm of prostate, ICD-10 diagnosis code C51 in death records) in a general population screening setting (such as in conjunction with a prostate-specific antigen (PSA) test), when the biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population.

Figure 3 shows the relation of baseline concentrations of 9 blood biomarkers to breast cancer survival, analysed as breast cancer mortality (malignant neoplasm of breast, ICD-10 diagnosis code C50 in death records) among females with an already diagnosed prevalent breast cancer at the time of the blood sampling (ICD-10 diagnosis code C50 in hospital records), when the biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population.

Figure 4 shows the relation of baseline concentrations of 9 blood biomarkers to prostate cancer survival, analysed as prostate cancer mortality (malignant neoplasm of prostate, ICD-10 diagnosis code C61 in death records) among males with an already diagnosed prevalent prostate cancer at the time of the blood sampling (ICD-10 diagnosis code C61 in hospital records), when the biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population.

Figure 5a shows an intended use case for a multibiomarker score to predict breast cancer mortality risk in a general population screening setting (such as along with a mammography test).

5 **Figure 5b** shows that the prediction of the risk for breast cancer mortality works effectively for people with different demographics and risk factor profiles, with substantially stronger results for short term risk prediction.

10 **Figure 6a** shows an intended use case for a multibiomarker score to predict prostate cancer mortality risk in a general population screening setting (such as along with a prostate-specific antigen (PSA) test).

15 **Figure 6b** shows that the prediction of the risk for prostate cancer mortality works effectively for people with different demographics and risk factor profiles, with substantially stronger results for short term risk prediction.

20

DETAILED DESCRIPTION

A method for determining whether a subject is at risk of dying from breast cancer is disclosed.

25 The method may comprise determining in a biological sample obtained from the subject a quantitative value of at least one biomarker of the following:

- glycoprotein acetyls,
- a ratio of docosahexaenoic acid to total fatty acids,
- 30 - a ratio of linoleic acid to total fatty acids,
- a ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids,
- 35 - a ratio of omega-3 fatty acids to total fatty acids,

- a ratio of omega-6 fatty acids to total fatty acids,
- a ratio of saturated fatty acids to total fatty acids,
- 5 - fatty acid degree of unsaturation,
- histidine; and

and comparing the quantitative value(s) of the at least one biomarker to a control sample or to a control value;

10 wherein an increase or a decrease in the quantitative value(s) of the at least one biomarker, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of dying from breast cancer.

15 The method is a method for determining whether a subject is at risk of dying from breast cancer. In such an embodiment, the subject may be a woman, i.e. female.

The control sample or control value may 20 represent women without a known breast cancer diagnosis, such as in a population screening setting (such as along with a mammography). Alternatively or additionally, the control sample or control value may represent women already diagnosed with breast cancer. This may allow 25 predicting the prognosis of the cancer.

The subject is a woman without known diagnosis of breast cancer.

The subject may be a woman in a general population screening setting, such as a mammography.

30 Various blood biomarkers may be useful for predicting whether an individual person is at an elevated risk of dying from breast cancer. Such biomarkers may be measured from biological samples, for example from blood samples or related biological fluids.

35 Biomarkers predictive of the risk for dying from breast cancer could help to enable more effective screening and better targeted preventative treatment,

and/or guide for decision on, for instance, prophylactic mastectomy.

The method comprises determining a quantitative value of glycoprotein acetyls.

5 In an embodiment, the method comprises determining a quantitative value of the ratio of docosaheptaenoic acid to total fatty acids.

10 In an embodiment, the method comprises determining a quantitative value of the ratio of linoleic acid to total fatty acids.

In an embodiment, the method comprises determining a quantitative value of the ratio of monounsaturated fatty acids and/or oleic acid to total fatty acids.

15 In an embodiment, the method comprises determining a quantitative value of the ratio of omega-3 fatty acids to total fatty acids.

20 In an embodiment, the method comprises determining a quantitative value of the ratio of omega-6 fatty acids to total fatty acids.

In an embodiment, the method comprises determining a quantitative value of the ratio of saturated fatty acids to total fatty acids.

25 In an embodiment, the method comprises determining a quantitative value of fatty acid degree of unsaturation.

In an embodiment, the method comprises determining a quantitative value of histidine.

30 The metabolic biomarker(s) described in this specification have been found to be significantly different, i.e. their quantitative values (such as amount and/or concentration) have been found to be significantly higher or lower, for subjects who later died from breast cancer. The biomarkers may be detected
35 and quantified from blood, serum, or plasma, dry blood spots, or other suitable biological sample, and may be used to predict the prognosis and/or survival, i.e. to

determine the risk of dying from breast cancer, either alone or in combination with other biomarkers.

Furthermore, the biomarker(s) may significantly improve the possibility of identifying
5 subjects at risk of dying from breast cancer, even in combination with and/or when accounting for established factors that may currently be used for prediction of the prognosis and/or survival, such as age, family history, smoking status, body mass index (BMI), existing
10 comorbidities, inflammatory protein biomarkers such as C-reactive protein and/or albumin, stage of the cancer and metastasis, genetic variants including BRCA1 and BRCA2 alleles, tumor size, tumor grade, certain traits of the cancer cells and/or response to the treatment.
15 The biomarkers described in this specification, alone or as a risk score (such as a multi-biomarker score), hazard ratio, odds ratio, and/or predicted absolute or relative risk, or in combination with other risk factors and tests, such as mammography, may improve the
20 prediction accuracy. This may include improving prediction accuracy by complementing the predictive information from other risk factors, or by replacing the need for other analyses. The biomarkers or the risk score, hazard ratio, odds ratio, and/or predicted
25 absolute or relative risk according to one or more embodiments described in this specification may thus allow for efficiently assessing the risk for dying from breast cancer, also in conditions in which other risk factor measures are not as feasible.

30 In an embodiment, the method is a method for determining whether the subject is at risk of dying from breast cancer. The method may be applied, for instance, in general population screening settings, such as along with a mammography, or among individuals already
35 diagnosed with breast cancer to predict the prognosis of the cancer.

In an embodiment, the method is a method for determining whether the subject is at risk of dying from breast cancer within (the following) three years.

The method may comprise determining in the
5 biological sample quantitative values of a plurality of the biomarkers, such as two, three, four, five or more biomarkers. For example, the plurality of the biomarkers may comprise 2, 3, 4, 5, 6, 7, 8, 9, or 10 (i.e. at least 2, at least 3, at least 4, at least 5, at least
10 6, at least 7, at least 8, at least 9, or all of the biomarkers). The term "plurality of the biomarkers" may thus, within this specification, be understood as referring to any number (above one) of the biomarkers. The term "plurality of the biomarkers" may thus be
15 understood as referring to any number (above one) and/or combination or subset of the biomarkers described in this specification. Determining the plurality of the biomarkers may increase the accuracy of the prediction of whether the subject is at risk of dying from breast
20 cancer. In general, it may be that the higher the number of the biomarkers, the more accurate or predictive the method. However, even a single biomarker described in this specification may allow for or assist in determining whether the subject is at risk of dying from
25 breast cancer. The plurality of the biomarkers may be measured from the same biological sample or from separate biological samples and using the same analytical method or different analytical methods. In an embodiment, the plurality of biomarkers may be a
30 panel of a plurality of biomarkers.

In the context of this specification, the wording "comparing the quantitative value(s) of the biomarker(s) to a control sample or to a control value(s)" may be understood, as a skilled person would,
35 as referring to comparing the quantitative value or values of the biomarker or biomarkers, to a control sample or to a control value(s) either individually or

as a plurality of biomarkers (e.g. when a risk score is calculated from the quantitative values of a plurality of biomarkers), depending e.g. on whether the quantitative value of a single (individual) biomarker or the quantitative values of a plurality of biomarkers are determined. The control sample or control value may represent women without known breast cancer diagnosis, for example as in population screening settings (such as along with a mammography), or women already diagnosed with breast cancer, to predict the prognosis of the cancer.

In an embodiment, the method may comprise determining in the biological sample obtained from the subject a quantitative value of the following biomarkers:

- glycoprotein acetyls,
 - a ratio of docosahexaenoic acid to total fatty acids,
 - a ratio of linoleic acid to total fatty acids,
 - a ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids,
 - a ratio of omega-3 fatty acids to total fatty acids,
 - a ratio of omega-6 fatty acids to total fatty acids,
 - a ratio of saturated fatty acids to total fatty acids,
 - fatty acid degree of unsaturation,
 - histidine; and
- comparing the quantitative value(s) of the biomarkers to a control sample or to a control value(s); wherein an increase or a decrease in the quantitative value(s) of the biomarkers, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of dying from breast cancer.

The at least one biomarker comprises or is glycoprotein acetyls. The method may further comprise determining a quantitative value of at least one of the other biomarkers described in this specification.

5 The subject may be human. The human may be healthy or have an existing disease, such as an existing cancer. Specifically, the human may have an already existing form of breast cancer, and the risk for developing more severe forms of and/or dying from these
10 cancers may be determined and/or calculated.

 In the context of this specification, the term "biomarker" may refer to a biomarker, for example a chemical or a molecular marker, that may be found to be associated with a severe course of a disease or the risk
15 of dying from a disease. It does not necessarily refer to a biomarker that would be statistically fully validated as having a specific effectiveness in a clinical setting. The biomarker may be a metabolite, a compound, a lipid, a protein, a moiety, a functional
20 group, a composition, a combination of two or more metabolites and/or compounds, a (measurable or measured) quantity thereof, a ratio or other value derived thereof, or in principle any measurement reflecting a chemical and/or biological component that may be found
25 to be associated with dying from a disease. The biomarkers and any combinations thereof, optionally in combination with further analyses and/or measures, may be used to measure a biological process indicative of the risk for dying from breast cancer.

30 In the context of this specification, the term "dying from breast cancer" may be understood as referring to death caused by breast cancer. In other words, the subject having an increased risk of dying from breast cancer may be understood as the subject
35 having an increased risk of developing a fatal breast cancer, i.e. death from or caused by the breast cancer, e.g. either among individuals without known diagnosis

of these cancers (as in population screening settings) to reflect the risk of developing a fatal form of these cancers later in life or among individuals with an existing diagnosis of these cancers to reflect the survival after diagnosis.

The subject is a subject, such as a woman, without known diagnosis of breast cancer. The subject may, additionally or alternatively, be a woman in a general population screening setting, such as a mammography.

Breast and/or prostate cancers described in this specification may be classified as follows. "ICD-10" may be understood as referring to the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) - WHO Version for 2019. Similar cancers classified or diagnosed by other disease classification on systems than ICD-10, such as ICD-9 or ICD-11, may also apply.

Breast and prostate cancers may be understood as referring to diseases classified within the 3-character ICD-10 diagnoses for breast and prostate cancers (C50: malignant neoplasm of breast, C61: malignant neoplasm of prostate).

In an embodiment, breast cancer is a specific disease, such as defined by ICD-10 3-character code diagnoses.

In an embodiment, breast cancer is the following ICD-10 3-character diagnosis:

- C50: Malignant neoplasm of breast

In an embodiment, the death from breast cancer may comprise or be death from breast cancer, such as a death from a disease denoted by any one of the ICD-10 codes listed above.

In an embodiment, the breast cancer comprises or is malignant neoplasm of breast.

The method may further comprise determining whether the subject is at risk of dying from breast

cancer using a risk score, hazard ratio, odds ratio, and/or predicted absolute or relative risk calculated on the basis of the quantitative value(s) of the at least one biomarker or of the plurality of the
5 biomarkers.

An increase or a decrease in the risk score, hazard ratio, and/or predicted absolute risk and/or relative risk may be indicative of the subject having an increased risk of dying from breast cancer. The
10 prediction may be applied, for instance, among women without known breast cancer diagnosis, as in a general population screening setting, such as along with a mammography, or among individuals already diagnosed with breast cancer to predict the survival.

The risk score and/or hazard ratio and/or predicted absolute risk or relative risk may be calculated based on any plurality, combination or subset of biomarkers described in this specification.
15

The risk score and/or hazard ratio and/or predicted absolute risk or relative risk may be calculated e.g. as shown in the Examples below. For example, the plurality of biomarkers measured using a suitable method, for example with nuclear magnetic resonance (NMR) spectroscopy, may be combined using
20 regression algorithms and multivariate analyses and/or using machine learning analysis. Before regression analysis or machine learning, any missing values in the biomarkers may be imputed with the mean value of each biomarker for the dataset. A number of biomarkers, for
25 example five, that may be considered most associated with death from the disease may be selected for the use in the prediction model. Other modelling approaches may be used to calculate a risk score and/or hazard ratio and/or predicted absolute risk or relative risk based
30 on a combination or subset of individual biomarkers, i.e. a plurality of the biomarkers.
35

The risk score may be calculated e.g. as a weighted sum of individual biomarkers, i.e. a plurality of the biomarkers. The weighted sum may be e.g. in the form of a multi-biomarker score defined as $\sum_i [\beta_i * c_i] + \beta_0$; where i is the index of summation over individual biomarkers, β_i is the weighted coefficient attributed to biomarker i , c_i is the blood concentration of biomarker i , and β_0 is an intercept term.

For example, the risk score can be defined as:
 $\beta_1 * \text{concentration}(\text{glycoprotein acetyls}) + \beta_2 * \text{concentration}(\text{monounsaturated fatty acid ratio to total fatty acids}) + \beta_3 * \text{concentration}(\text{histidine}) + \beta_0$, where β_1 , β_2 , β_3 are multipliers for each biomarker according to the association magnitude with risk of dying from breast cancer and β_0 is an intercept term. As a skilled person will understand, the biomarkers mentioned in this example may be replaced by any other biomarker(s) described in this specification. In general, the more biomarkers are included in the risk score, the stronger the predictive performance may become. When additional biomarkers are included in the risk score, the β_i weights may change for all biomarkers according to the optimal combination for the prediction of death from breast cancer.

The risk score, hazard ratio, odds ratio, and/or predicted relative risk and/or absolute risk may be calculated on the basis of at least one further measure, for example a characteristic of the subject. Such characteristics may be determined prior to, simultaneously, or after the biological sample is obtained from the subject. As a skilled person will understand, some of the characteristics may be information collected e.g. using a questionnaire or clinical data collected earlier. Some of the characteristics may be determined (or may have been determined) by biochemical or clinical diagnostic measurements and/or medical diagnosis. Such

characteristics could include, for example, one or more of age, height, weight, body mass index, race or ethnic group, smoking, family history of cancers and/or information from mammography.

5 The risk prediction for death from breast cancer guided based on one or more of the biomarkers can be used to guide treatment and preventative efforts for a severe course of the cancer, such as alcohol and smoking awareness, healthy diet, physical activity,
10 clinical screening frequency and/or treatment decisions. For example, the information of the risk for death from breast cancer can be used for guiding frequency of cancer screenings, decision on prophylactic mastectomy and/or treatment with, for instance, aspirin
15 and other non-steroidal anti-inflammatory drugs (NSAIDs), and/or selective estrogen receptor modulators (SERMs), such as tamoxifen and/or raloxifene.

 In the context of this specification, the term "glycoprotein acetyls", "glycoprotein acetylation", or
20 "GlycA" may refer to a nuclear magnetic resonance spectroscopy (NMR) signal that represents the abundance of circulating glycosylated proteins, i.e. N-acetylated glycoproteins. Glycoprotein acetyls may include signals from a plurality of different glycoproteins, including
25 e.g. alpha-1-acid glycoprotein, alpha-1 antitrypsin, haptoglobin, transferrin, and/or alpha-1 antichymotrypsin. In the scientific literature on cardiometabolic biomarkers, the terms "glycoprotein acetyls" or "GlycA" may commonly refer to the NMR signal
30 of circulating glycosylated proteins (e.g. Ritchie et al, Cell Systems 2015 1(4):293-301; Connelly et al, J Transl Med. 2017;15(1):219). Glycoprotein acetyls and a method for measuring them is described e.g. in Kettunen et al., 2018, Circ Genom Precis Med. 11:e002234 and Soininen et
35 al., 2009, Analyst 134, 1781-1785. There may be benefits of using the NMR signal of glycoprotein acetyls for risk prediction above measurement of the individual proteins

contributing to the NMR signal, for instance better analytical accuracy and stability over time, as well as lower costs of the measurement and the possibility to measure the NMR signal simultaneously with many other
5 biomarkers.

In the context of this specification, the term "omega-3 fatty acids" may refer to total omega-3 fatty acids, i.e. the total omega-3 fatty acid amounts and/or concentrations, i.e. the sum of different omega-3 fatty
10 acids. Omega-3 fatty acids are polyunsaturated fatty acids. In omega-3 fatty acids, the last double bond in the fatty acid chain is the third bond counting from the methyl end. Docosahexaenoic acid is an example of an omega-3 fatty acid.

In the context of this specification, the term "omega-6 fatty acids" may refer to total omega-6 fatty acids, i.e. the total omega-6 fatty acid amounts and/or concentrations, i.e. the sum of the amounts and/or concentrations of different omega-6 fatty acids. Omega-
15 6 fatty acids are polyunsaturated fatty acids. In omega-6 fatty acids, the last double bond in the fatty acid chain is the sixth bond counting from the methyl end.

In one embodiment, the omega-6 fatty acid may be linoleic acid. Linoleic acid (18:2 ω -6) is the most
25 abundant type of omega-6 fatty acids, and may therefore be considered as a good approximation for total omega-6 fatty acids for risk prediction of dying from breast cancer.

In the context of this specification, the term
30 "monounsaturated fatty acids" (MUFAs) may refer to total monounsaturated fatty acids, i.e. the total MUFA amounts and/or concentrations. Monounsaturated fatty acids may, alternatively, refer to oleic acid, which is the most abundant monounsaturated fatty acid in human serum.
35 Monounsaturated fatty acids have one double bond in their fatty acid chain. The monounsaturated fatty acids may include omega-9 and omega-7 fatty acids. Oleic acid

(18:1 ω -9), palmitoleic acid (16:1 ω -7) and cis-vaccenic acid (18:1 ω -7) are examples of common monounsaturated fatty acids in human serum.

5 In one embodiment, the monounsaturated fatty acid may be oleic acid. Oleic acid is the most abundant monounsaturated fatty acid, and may therefore be considered as a good approximation for total monounsaturated fatty acids for risk prediction of dying from breast cancer.

10 In the context of this specification, the term "saturated fatty acids" (SFAs) may refer to total saturated fatty acids. Saturated fatty acids may be or comprise fatty acids which have no double bonds in their structure. Palmitic acid (16:0) and stearic acid (18:0)
15 are examples of abundant SFAs in human serum.

For all fatty acid measures, including omega-6, docosahexaenoic acid, linoleic acid, monounsaturated fatty acids and/or saturated fatty acids, the fatty acid measures may include blood (or serum/plasma) free fatty
20 acids, bound fatty acids and esterified fatty acids. Esterified fatty acids may, for example, be esterified to glycerol as in triglycerides, diglycerides, monoglycerides, or phosphoglycerides, or to cholesterol as in cholesterol esters.

25 In the context of this specification, the term "fatty acid degree of unsaturation" or "unsaturation" may be understood as referring to the number of double bonds in total fatty acids, for example the average number of double bonds in total fatty acids.

30 In the context of this specification, the term "histidine" may refer to the histidine amino acid, for example in blood, plasma or serum or related biofluids.

In the context of this specification, the term "quantitative value" may refer to any quantitative value
35 characterizing the amount and/or concentration of a biomarker. For example, it may be the amount or concentration of the biomarker in the biological sample,

or it may be a signal derived from nuclear magnetic resonance spectroscopy (NMR) or other method suitable for detecting the biomarker in a quantitative manner. Such a signal may be indicative of or may correlate with the amount or concentration of the biomarker. It may also be a quantitative value calculated from one or more signals derived from NMR measurements or from other measurements. Quantitative values may, additionally or alternatively, be measured using a variety of techniques. Such methods may include mass spectrometry (MS), gas chromatography combined with MS, high performance liquid chromatography alone or combined with MS, immunoturbidimetric measurements, ultracentrifugation, ion mobility, enzymatic analyses, colorimetric or fluorometric analyses, immunoblot analysis, immunohistochemical methods (e.g. *in situ* methods based on antibody detection of metabolites), and immunoassays (e.g. ELISA). Examples of various methods are set out below. The method used to determine the quantitative value(s) in the subject may be the same method that is used to determine the quantitative value(s) in a control subject/control subjects or in a control sample/control samples.

The quantitative value, or the initial quantitative value, of the at least one biomarker, or the plurality of the biomarkers, may be measured using nuclear magnetic resonance (NMR) spectroscopy, for example ^1H -NMR. The at least one additional biomarker, or the plurality of the additional biomarkers, may also be measured using NMR. NMR may provide a particularly efficient and fast way to measure biomarkers, including a large number of biomarkers simultaneously, and can provide quantitative values for them. NMR also typically requires very little sample pre-treatment or preparation. The biomarkers measured with NMR can effectively be measured for large amounts of samples using an assay for blood (serum or plasma) NMR

metabolomics previously published by Soininen et al., 2015, *Circulation: Cardiovascular Genetics* 8, 212-206 (DOI: 10.1161/CIRCGENETICS.114.000216); Soininen et al., 2009, *Analyst* 134, 1781-1785; and Würtz et al., 5 2017, *American Journal of Epidemiology* 186 (9), 1084-1096 (DOI: 10.1093/aje/kwx016). This provides data on 250 biomarkers per sample as described in detail in the above scientific papers.

In an embodiment, the (initial) quantitative 10 value(s) of the at least one biomarker is/are measured using nuclear magnetic resonance spectroscopy.

However, quantitative values for various biomarkers described in this specification may also be performed by techniques other than NMR. For example, 15 mass spectrometry (MS), enzymatic methods, antibody-based detection methods, or other biochemical or chemical methods may be contemplated, depending on the biomarker.

20 E.g. monounsaturated fatty acids, saturated fatty acids, and omega-6 fatty acids can be quantified (i.e. their quantitative values may be determined) by serum total fatty acid composition using gas chromatography (for example, as described in Jula et 25 al., 2005, *Arterioscler Thromb Vasc Biol* 25, 2152-2159).

In the context of this specification, the term "sample" or "biological sample" may refer to any biological sample obtained from a subject or a group or population of subjects. The sample may be fresh, frozen, 30 or dry.

The biological sample may comprise or be, for example, a blood sample, a plasma sample, a serum sample, or a sample or fraction derived therefrom. The biological sample may be, for example, a fasting blood 35 sample, a fasting plasma sample, a fasting serum sample, or a fraction obtainable therefrom. However, the biological sample does not necessarily have to be a

fasting sample. The blood sample may be a venous blood sample.

The blood sample may be a dry blood sample. The dry blood sample may be a dried whole blood sample, a dried plasma sample, a dried serum sample, or a dried sample derived therefrom.

The biological sample may be obtained from the subject prior to determining the quantitative value of the at least one biomarker. Taking a blood sample or a tissue sample of a subject or patient is a part of normal clinical practice. The collected blood or tissue sample can be prepared and serum or plasma can be separated using techniques well known to a skilled person. Methods for separating one or more fractions from biological samples, such as blood samples or tissue samples, are also available to a skilled person. The term "fraction" may, in the context of this specification, also refer to a portion or a component of the biological sample separated according to one or more physical properties, for instance solubility, hydrophilicity or hydrophobicity, density, or molecular size.

In the context of this specification, the term "control sample" may refer to a sample obtained from a subject and known not to suffer from the disease or condition or not being at risk of having or developing the disease or condition. In an embodiment, the control sample may be a biological sample from a subject, such as a woman, or a generalized population of women without known positive diagnosis of breast cancer. In other words, the control sample may be a biological sample from a woman or a generalized population of women who has/have not been (previously) diagnosed with breast cancer. Alternatively, in an embodiment, the control population may be a sample of women already diagnosed with breast cancer. It may be compared to samples developing and/or dying from these cancers to predict the mortality risk. The control sample may be matched.

The term "control value" may be understood as a value obtainable from the control sample and/or a quantitative value derivable therefrom. In an embodiment, the control value may be a value derivable from one or more women
5 or a generalized population of women without known positive diagnosis of breast cancer, respectively. For example, it may be possible to calculate a threshold value from control samples and/or control values, above or below which the risk of dying from the cancer is
10 elevated. In other words, a value higher or lower (depending on the biomarker, risk score, hazard ratio, and/or predicted absolute risk or relative risk) than the threshold value may be indicative of the subject having an increased risk of dying from the cancer. The
15 risk score, hazard ratio, odds ratio, and/or predicted absolute or relative risk may be used alone or in combination with other risk factors and measures, such as information from mammography and/or family history.

An increase or a decrease in the quantitative
20 value(s) of the at least one biomarker, or the plurality of the biomarkers, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from the cancer. Whether an increase or a decrease is indicative of the
25 subject having an increased risk of dying from the cancer, may depend on the biomarker.

A 1.2-fold, 1.5-fold, or for example 2-fold, or 3-fold, increase or a decrease in the quantitative
30 value(s) of the at least one biomarker (or in an individual biomarker of the plurality of the biomarkers) when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from the cancer.

In an embodiment, an increase in the
35 quantitative value of glycoprotein acetyls, when compared to the control sample or to the control value,

may be indicative of the subject having an increased risk of dying from breast cancer.

5 In an embodiment, a decrease in the quantitative value of docosahexaenoic acid and/or the ratio of docosahexaenoic acid to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

10 In an embodiment, a decrease in the quantitative value of linoleic acid and/or the ratio linoleic acid to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

15 In an embodiment, an increase in the quantitative value of monounsaturated fatty acids and/or oleic acid and/or the ratio of monounsaturated fatty acids and/or oleic acid to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

20 In an embodiment, a decrease in the quantitative value of omega-3 fatty acids and/or the ratio of omega-3 fatty acids to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

25 In an embodiment, a decrease in the quantitative value of omega-6 fatty acids and/or the ratio of omega-6 fatty acids to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

30 In an embodiment, an increase in the quantitative value of saturated fatty acids and/or the ratio of saturated fatty acids to total fatty acids, when compared to the control sample or to the control

35

value, may be indicative of the subject having an increased risk of dying from breast cancer.

5 In an embodiment, a decrease in the quantitative value of fatty acid degree of unsaturation, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

10 In an embodiment, a decrease in the quantitative value of histidine, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

15 In an embodiment, a risk score defined as $\beta_0 + \beta_1^*$ concentration (glycoprotein acetyls) + β_2^* concentration (fatty acid measure), where β_0 is an intercept term, β_1 is the weighted coefficient attributed to the concentration of glycoprotein acetyls, β_2 is the weighted coefficient attributed to the fatty acid measure, may be indicative of the subject having an increased risk of dying from breast cancer. The fatty acid measure may be one or more of the following fatty acid measures: the ratio of docosahexaenoic acid to total fatty acids, the ratio of linoleic acid to total fatty acids, the ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids, the ratio of omega-3 fatty acids to total fatty acids, the ratio of omega-6 fatty acids to total fatty acids, the ratio of saturated fatty acids to total fatty acids, fatty acid degree of unsaturation.

30 In an embodiment, a risk score defined as $\beta_0 + \beta_1^*$ concentration (glycoprotein acetyls) + β_2^* concentration (histidine) + β_3^* concentration (fatty acid measure), where β_0 is an intercept term, β_1 is the weighted coefficient attributed to the concentration of glycoprotein acetyls, β_2 is the weighted coefficient attributed to the concentration of histidine, and β_3 is the weighted coefficient attributed to the concentration

35

of the fatty acid measure may be indicative of the subject having an increased risk of dying from breast cancer. The fatty acid measure may be one or more of the following fatty measures: the ratio of docosahexaenoic acid to total fatty acids, the ratio of linoleic acid to total fatty acids, the ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids, the ratio of omega-3 fatty acids to total fatty acids, the ratio of omega-6 fatty acids to total fatty acids, the ratio of saturated fatty acids to total fatty acids, fatty acid degree of unsaturation.

The term "combination" may, at least in some embodiments, be understood such that the method comprises using a risk score, hazard ratio, odds ratio, and/or predicted absolute risk or relative risk calculated on the basis of the quantitative value(s) of the biomarkers. For example, if quantitative values of both glycoprotein acetyls and a fatty acid measure are determined, the quantitative values of both biomarkers may be compared to the control sample or the control value separately, or a risk score, hazard ratio, odds ratio, and/or predicted absolute risk or relative risk calculated on the basis of the quantitative value(s) of both the biomarkers, and the risk score, hazard ratio, odds ratio, and/or predicted absolute risk or relative risk may be compared to the control sample or the control value.

In an embodiment, the method may comprise determining in the biological sample obtained from the subject a quantitative value of the following biomarkers:

- glycoprotein acetyls,
- at least one fatty acid measure(s) of the following: the ratio of docosahexaenoic acid to total fatty acids, the ratio of linoleic acid to total fatty acids, the ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids, the ratio of omega-

3 fatty acids to total fatty acids, the ratio of omega-6 fatty acids to total fatty acids, the ratio of saturated fatty acids to total fatty acids, fatty acid degree of unsaturation; and

5 comparing the quantitative value(s) of the biomarkers and/or a combination thereof to a control sample or to a control value(s);

wherein an increase or a decrease in the quantitative value(s) of the biomarkers and/or the combination thereof, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of dying from breast cancer. An increase in the quantitative value of glycoprotein acetyls and a decrease in the quantitative value of the ratio of docosahexaenoic and/or linoleic acid and/or omega-3 fatty acids and/or omega-6 fatty acids to total fatty acids and/or fatty acid degree of unsaturation, and/or an increase in the quantitative value of the ratio of monounsaturated fatty acids and/or saturated fatty acids to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

In an embodiment, the method may comprise determining in the biological sample obtained from the subject a quantitative value of the following biomarkers:

- glycoprotein acetyls,
- histidine,
- 30 - at least one fatty acid measure(s) of the following: the ratio of docosahexaenoic acid to total fatty acids, the ratio of linoleic acid to total fatty acids, the ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids, the ratio of omega-3 fatty acids to total fatty acids, the ratio of omega-6 fatty acids to total fatty acids, the ratio of

saturated fatty acids to total fatty acids, fatty acid degree of unsaturation; and

5 comparing the quantitative value(s) of the biomarkers and/or a combination thereof to a control sample or to a control value(s);

10 wherein an increase or a decrease in the quantitative value(s) of the biomarkers and/or the combination thereof, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of dying from breast cancer. An increase in the quantitative value of glycoprotein acetyls, a decrease in the quantitative value of histidine, and a decrease in the quantitative value of the ratio of docosahexaenoic and/or linoleic acid and/or omega-3 fatty acids and/or omega-6 fatty acids to total fatty acids and/or fatty acid degree of unsaturation, and/or an increase in the quantitative value of the ratio of monounsaturated fatty acids and/or saturated fatty acids and/or to total fatty acids, when compared to the control sample or to the control value, may be indicative of the subject having an increased risk of dying from breast cancer.

EXAMPLES

25 Reference will now be made in detail to various embodiments, an example of which is illustrated in the accompanying drawings. The description below discloses some embodiments in such a detail that a person skilled in the art is able to utilize the embodiments based on the disclosure. Not all steps or features of the 30 embodiments are discussed in detail, as many of the steps or features will be obvious for the person skilled in the art based on this specification.

35 Abbreviations used in the Figures:

DHA %: Ratio of docosahexaenoic acid to total fatty acids

LA%: Ratio of linoleic acid to total fatty acids
 MUFA %: Ratio of monounsaturated fatty acids to total fatty acids
 5 Omega-3 %: Ratio of omega-3 fatty acids to total fatty acids
 Omega-6 %: Ratio of omega-6 fatty acids to total fatty acids
 10 SFA %: Ratio of saturated fatty acids to total fatty acids
 Unsaturation: Fatty acid degree of unsaturation
 CI: confidence interval
 SD: standard deviation
 15 BMI: Body mass index

EXAMPLE 1

Biomarker measures quantified by NMR spectroscopy were investigated as to whether they could
 20 be predictive of dying from breast cancer or prostate cancer. All analyses were conducted based on the UK Biobank, with approximately 115 000 study participants with blood biomarker data from NMR spectroscopy available.

25 Study population

Details of the design of the UK Biobank have been reported by Sudlow et al 2015, PLoS Med. 2015;12(3):e1001779. Briefly, UK Biobank recruited 502
 30 639 participants aged 37-73 years in 22 assessment centres across the UK. All participants provided written informed consent and ethical approval was obtained from the North West Multi-Center Research Ethics Committee. Blood samples were drawn at baseline between 2007 and
 35 2010. No selection criteria were applied to the sampling.

Biomarker profiling

From the entire UK Biobank population, a random subset of baseline plasma samples from 118 466 individuals were measured using the Nightingale NMR biomarker platform (Nightingale Health Ltd, Finland).
5 This blood analysis method provides simultaneous quantification of many blood biomarkers, including lipoprotein lipids, circulating fatty acids, and various low-molecular weight metabolites including amino acids, ketone bodies and gluconeogenesis-related metabolites
10 in molar concentration units. Technical details and epidemiological applications have been reviewed (Soininen et al 2015, Circ Cardiovasc Genet; 2015;8:192-206; Würtz et al 2017, Am J Epidemiol 2017;186:1084-1096). Values outside four interquartile ranges from
15 median were considered as outliers and excluded.

Epidemiological analyses of biomarker relations with the risk of dying from breast cancer or prostate cancer

The blood biomarker associations with the risk
20 of dying from breast cancer or prostate cancer were conducted based on UK Biobank data. Analyses focused on the relation of the biomarkers to the mortality from breast cancer or prostate cancer after the blood samples were collected, to determine if the individual bi-
25 omarkers associate with the risk of dying from breast cancer or prostate cancer. Examples using multi-biomarker scores, in the form weighted sums of biomarkers, were also explored to demonstrate their effectiveness in various prediction settings.

30 Information on the disease and mortality events for all study participants were recorded from UK Hospital Episode Statistics data and death registries. Analyses were performed using two approaches: by analyzing the mortality risk among women and men
35 without known positive diagnosis of prostate or breast cancer, respectively, as in general population screening settings (such as along with a mammography or PSA

screening test, respectively), as well as among women or men already diagnosed with breast cancer or prostate cancer at the time of the blood sampling, respectively, to predict the prognosis of the cancer. Breast and prostate cancers analyzed here were defined according to the ICD-10 diagnosis codes C50 (malignant neoplasm of breast) and C61 (malignant neoplasm of prostate). The registry-based follow-up was from blood sampling in 2007-2010 through to 2020 (approximately 1 100 000 person-years).

For biomarker association testing, Cox proportional-hazard regression models adjusted for age and UK Biobank assessment centre were used. Results were plotted in magnitudes per standard deviation of each biomarker measure to allow direct comparison of association magnitudes.

Summary of results

Baseline characteristics of the study population for biomarker analyses are shown in Table 1.

Table 1: Clinical characteristics of study participants.

Total number of individuals with blood samples analysed	118 456
Study setting	Population sample of study volunteers from the UK
Percentage of women	54.1%
Age range (years)	39 - 71
Median age (years)	58
Median BMI (kg/m ²)	26.8
Follow-up time for disease events after blood sampling	10-14 years

Breast cancer (C50)	
Number of total breast cancer mortality events	311
Number of participants with prevalent breast cancer	1 536
Number of breast cancer mortality events among participants with prevalent breast cancer	130
Prostate cancer (C61)	
Number of prostate cancer mortality events	256
Number of participants with prevalent prostate cancer	584
Number of prostate cancer mortality events among participants with prevalent prostate cancer	55

Figure 1 shows the relation of baseline concentrations of 9 blood biomarkers to breast cancer mortality (ICD-10 diagnosis code C50 in death records) in a general population screening setting (such as along with a mammography test), when the biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population. The results are based on statistical analyses of 63 859 females from the UK Biobank, out of whom 311 individuals died from breast cancer during approximately 10 years of follow-up. The analyses were adjusted for age and UK Biobank assessment centre in Cox proportional-hazard regression models. In the forestplots, statistically significant results are displayed as filled points ($P < 0.001$ corresponding to multiple testing correction) and non-significant results are displayed as hollow points. These results demonstrate that the 9 individual biomarkers are predictive of the risk for dying from breast cancer in general population settings.

Figure 2 shows the relation of baseline concentrations of 9 blood biomarkers to prostate cancer mortality (ICD-10 diagnosis code C61 in death records) in a general population screening setting (such as along with a prostate-specific antigen (PSA) test), when the biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population. The results are based on statistical analyses of 54 042 males from the UK Biobank, out of whom 256 individuals died from prostate cancer during approximately 10 years of follow-up. The analyses were adjusted for age and UK Biobank assessment centre in Cox proportional-hazard regression models. In the forest-plots, statistically significant results are displayed as filled points ($P < 0.001$ corresponding to multiple testing correction) and non-significant results are displayed as hollow points. These results demonstrate that the 9 individual biomarkers are predictive of the risk for dying from prostate cancer in general population settings.

Figure 3 shows the relation of baseline concentrations of 9 blood biomarkers to breast cancer survival, here analysed as mortality from breast cancer (ICD-10 diagnosis code C50 in death records) among females with a prevalent breast cancer diagnosis at the time of the blood sampling (ICD-10 diagnosis code C50 in hospital records). The biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population. The results are based on statistical analyses of 1 526 females from the UK Biobank with a prevalent breast cancer at the time of the blood sampling, out of whom 130 individuals died from the breast cancer during approximately 10 years of follow-up. The analyses were adjusted for age and UK Biobank assessment centre in Cox proportional-hazard regression models. In the forestplots, statistically significant results are displayed as filled points

($P < 0.001$ corresponding to multiple testing correction) and non-significant results are displayed as hollow points. These results demonstrate that the 9 individual biomarkers are predictive of breast cancer survival among women with an existing form of breast cancer.

Figure 4 shows the relation of baseline concentrations of 9 blood biomarkers to prostate cancer survival, here analysed as mortality from prostate cancer (ICD-10 diagnosis code C61 in death records) among males with a prevalent prostate cancer diagnosis at the time of the blood sampling (ICD-10 diagnosis code C61 in hospital records). The biomarker concentrations are analysed in absolute concentrations scaled to standard deviations of the study population. The results are based on statistical analyses of 584 males from the UK Biobank with a prevalent prostate cancer at the time of the blood sampling, out of whom 55 individuals died from the prostate cancer during approximately 10 years of follow-up. The analyses were adjusted for age and UK Biobank assessment centre in Cox proportional-hazard regression models. In the forestplots, statistically significant results are displayed as filled points ($P < 0.001$ corresponding to multiple testing correction) and non-significant results are displayed as hollow points. These results demonstrate that the 9 individual biomarkers are predictive of prostate cancer survival among men with an existing form of prostate cancer.

Illustrations of intended use: biomarker scores for risk prediction of death from breast cancer or prostate cancer

For illustration of intended applications related to the prediction of dying from breast cancer or prostate cancer, further epidemiological analyses are illustrated below. Results are shown for a biomarker score combining the 9 biomarkers featured in Figures 1-4. Similar results, albeit slightly weaker, are obtained

with combinations of only two or three individual biomarkers.

Figure 5a shows the increase in the risk for breast cancer mortality (ICD-10 diagnosis code C50) in a general population screening setting (such as in conjunction with a mammography test) along with increasing levels of a biomarker risk score composed of a weighted sum of the 9 biomarkers. The risk increase is plotted in the form of Kaplan-Meier plots showing the cumulative mortality from breast cancer for selected quantiles of the plurality-biomarker-score.

Figure 5b shows the hazard ratios of the same plurality-biomarker score with breast cancer mortality risk when accounting for relevant risk factor characteristics of the study participants. The first panel demonstrates that the risk prediction works effectively for people at different ages at the time of blood sampling. The second panel shows that the magnitude of the hazard ratio is only modestly attenuated when accounting for body mass index and smoking status in the statistical modelling. The last panel demonstrates that the hazard ratios are substantially stronger for short term risk prediction.

Figure 6a shows the increase in the risk for prostate cancer mortality (ICD-10 diagnosis code C61) in a general population screening setting (such as in conjunction with a prostate-specific antigen (PSA) test) along with increasing levels of a biomarker risk score composed of a weighted sum of the 9 biomarkers. The risk increase is plotted in the form of Kaplan-Meier plots showing the cumulative mortality from prostate cancer for selected quantiles of the plurality-biomarker-score.

Figure 6b shows the hazard ratios of the same plurality-biomarker score with prostate cancer mortality risk when accounting for relevant risk factor characteristics of the study participants. The first panel

demonstrates that the risk prediction works effectively for people at different ages at the time of blood sampling. The second panel shows that the magnitude of the hazard ratio is only modestly attenuated when accounting
5 for body mass index and smoking status in the statistical modelling. The last panel demonstrates that the hazard ratios are substantially stronger for short term risk prediction.

10 It is obvious to a person skilled in the art that with the advancement of technology, the basic idea may be implemented in various ways. The embodiments are thus not limited to the examples described above; instead they may vary within the scope of the claims.

15 The embodiments described hereinbefore may be used in any combination with each other. Several of the embodiments may be combined together to form a further embodiment. A method disclosed herein may comprise at least one of the embodiments described hereinbefore. It
20 will be understood that the benefits and advantages described above may relate to one embodiment or may relate to several embodiments. The embodiments are not limited to those that solve any or all of the stated problems or those that have any or all of the stated
25 benefits and advantages. It will further be understood that reference to 'an' item refers to one or more of those items. The term "comprising" is used in this specification to mean including the feature(s) or act(s) followed thereafter, without excluding the presence of
30 one or more additional features or acts.

CLAIMS

1. A method for determining whether a subject is at risk of dying from breast cancer;

5 wherein the method comprises determining in a biological sample obtained from the subject a quantitative value of at least one biomarker of the following in the biological sample:

- 10 - glycoprotein acetyls,
- a ratio of docosahexaenoic acid to total fatty acids,
- a ratio of linoleic acid to total fatty acids,
- a ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids,
- 15 - a ratio of omega-3 fatty acids to total fatty acids,
- a ratio of omega-6 fatty acids to total fatty acids,
- a ratio of saturated fatty acids to total fatty acids,
- 20 - fatty acid degree of unsaturation,
- histidine; and

25 comparing the quantitative value(s) of the at least one biomarker to a control sample or to a control value;

30 wherein an increase or a decrease in the quantitative value(s) of the at least one biomarker, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of dying from breast cancer;

35 wherein the at least one biomarker comprises or is glycoprotein acetyls, wherein glycoprotein acetyls refers to a nuclear magnetic resonance spectroscopy signal that represents the abundance of circulating glycosylated proteins, and

wherein the subject is a woman without known diagnosis of breast cancer.

2. The method according to claim 1, wherein the method comprises determining in the biological sample quantitative values of a plurality of biomarkers, such as two, three, four, five or more biomarkers.

5 3. The method according to claim 1 or 2, wherein the method comprises determining in the biological sample obtained from the subject a quantitative value of the following biomarkers:

 - glycoprotein acetyls,
10 - at least one fatty acid measure(s) of the following: the ratio of docosahexaenoic acid to total fatty acids, the ratio of linoleic acid to total fatty acids, the ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids, the ratio of omega-
15 3 fatty acids to total fatty acids, the ratio of omega-6 fatty acids to total fatty acids, the ratio of saturated fatty acids to total fatty acids, fatty acid degree of unsaturation; and

 comparing the quantitative value(s) of the
20 biomarkers to a control sample or to a control value(s); wherein an increase or a decrease in the quantitative value(s) of the biomarkers, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of
25 dying from breast cancer.

4. The method according to any one of claims 1 - 3, wherein the method comprises determining in the biological sample obtained from the subject a quantitative value of the following biomarkers:

30 - glycoprotein acetyls,
 - histidine,
 - at least one fatty acid measure(s) of the following: the ratio of docosahexaenoic acid to total
35 fatty acids, the ratio of linoleic acid to total fatty acids, the ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids, the ratio of omega-3 fatty acids to total fatty acids, the ratio of omega-

6 fatty acids to total fatty acids, the ratio of saturated fatty acids to total fatty acids, fatty acid degree of unsaturation; and

5 comparing the quantitative value(s) of the biomarkers to a control sample or to a control value(s); wherein an increase or a decrease in the quantitative value(s) of the biomarkers, when compared to the control sample or to the control value, is/are indicative of the subject having an increased risk of
10 dying from breast cancer.

5. The method according to any one of claims 1 - 4, wherein the breast cancer comprises or is malignant neoplasm of breast (ICD-10 diagnosis code C50).

6. The method according to any one of claims 1
15 - 5, wherein the quantitative value of the at least one biomarker is/are measured using nuclear magnetic resonance spectroscopy.

7. The method according to any one of claims 1 - 6, wherein the method further comprises determining
20 whether the subject is at risk of dying from breast cancer using a risk score, hazard ratio, odds ratio, and/or predicted absolute risk or relative risk calculated on the basis of the quantitative value(s) of at least one biomarker or of the plurality of the
25 biomarkers.

8. The method according to any one of claims 1 - 7, wherein the control sample or control value represents women without known diagnosis of breast cancer or women already diagnosed with breast cancer.

30 9. The method according to any one of claims 1 - 8, wherein the subject is a woman in a general population screening setting, such as a mammography.

10. The method according to any one of claims 1 - 9, wherein the method comprises determining in the
35 biological sample obtained from the subject a quantitative value of the following biomarkers:

- glycoprotein acetyls,

- a ratio of docosahexaenoic acid to total fatty acids,
- a ratio of linoleic acid to total fatty acids,
- 5 - a ratio of monounsaturated fatty acids and/or of oleic acid to total fatty acids,
- a ratio of omega-3 fatty acids to total fatty acids,
- a ratio of omega-6 fatty acids to total fatty
- 10 acids,
- a ratio of saturated fatty acids to total fatty acids,
- fatty acid degree of unsaturation,
- histidine; and
- 15 comparing the quantitative value(s) of the biomarkers to a control sample or to a control value(s);
- wherein an increase or a decrease in the quantitative value(s) of the biomarkers, when compared to the control sample or to the control value, is/are
- 20 indicative of the subject having an increased risk of dying from breast cancer.

PATENTTIVAATIMUKSET

1. Menetelmä sen määrittämiseksi, onko kohteella riski kuolla rintasyöpään;

5 jossa menetelmä käsittää, että määritetään biologisesta näytteestä, joka saadaan kohteelta, seuraavista vähintään yhden biomarkkerin kvantitatiivinen arvo biologisessa näytteessä:

- glykoproteiiniasetyyliit,
- dokosaheksaeenihapon suhde
- 10 kokonaisrasvahappoihin,
- linolihapon suhde kokonaisrasvahappoihin,
- kertatyydyttymättömien rasvahappojen ja/tai oleiinihapon suhde kokonaisrasvahappoihin,
- omega-3-rasvahappojen suhde
- 15 kokonaisrasvahappoihin,
- omega-6-rasvahappojen suhde kokonaisrasvahappoihin,
- tyydyttyneiden rasvahappojen suhde kokonaisrasvahappoihin,
- 20 - rasvahappotyydyttymättömyysaste,
- histidiini; ja

verrataan vähintään yhden biomarkkerin kvantitatiivista arvoa tai kvantitatiivisia arvoja kontrollinäytteeseen tai kontrolliarvoon;

25 jossa vähintään yhden biomarkkerin kvantitatiivisen arvon tai kvantitatiivisten arvojen nousu tai lasku verrattuna kontrollinäytteeseen tai kontrolliarvoon osoittaa/osoittavat kohteen kohonnutta riskiä kuolla rintasyöpään;

30 jossa vähintään yksi biomarkkeri käsittää tai on glykoproteiiniasetyyliit, jossa glykoproteiiniasetyyliit viittaa ydinmagneettiseen resonanssispektroskopiasignaaliin, joka kuvaa kiertävien glykoituneiden proteiinien määrää, ja

35 jossa kohde on nainen, jolla ei ole tiedossa olevaa rintasyöpädiagnoosia.

2. Patenttivaatimuksen 1 mukainen menetelmä, jossa menetelmä käsittää, että määritetään biologisesta näytteestä useiden biomarkkerien, kuten kahden, kolmen, neljän, viiden tai useamman biomarkkerin, kvantitatiiviset arvot.

3. Patenttivaatimuksen 1 tai 2 mukainen menetelmä, jossa menetelmä käsittää, että määritetään biologisesta näytteestä, joka saadaan kohteelta, seuraavien biomarkkerien kvantitatiivinen arvo:

10 - glykoproteiiniasetyyliit,
 - vähintään yksi rasvahappomäärä (t)
 seuraavista: dokosaheksaeenihapon suhde
 kokonaisrasvahappoihin, linolihapon suhde
 kokonaisrasvahappoihin, kertatyydyttymättömien
 15 rasvahappojen ja/tai oleiinihapon suhde
 kokonaisrasvahappoihin, omega-3-rasvahappojen suhde
 kokonaisrasvahappoihin, omega-6-rasvahappojen suhde
 kokonaisrasvahappoihin, tyydyttyneiden rasvahappojen
 suhde kokonaisrasvahappoihin,
 20 rasvahappotyydyttymättömyysaste; ja

verrataan biomarkkerien kvantitatiivista arvoa tai kvantitatiivisia arvoja kontrollinäytteeseen tai kontrolliarvoon tai kontrolliarvoihin;

jossa biomarkkerien kvantitatiivisen arvon tai kvantitatiivisten arvojen nousu tai lasku verrattuna kontrollinäytteeseen tai kontrolliarvoon osoittaa/osoittavat kohteen kohonnutta riskiä kuolla rintasyöpään.

4. Jonkin patenttivaatimuksista 1 - 3 mukainen menetelmä, jossa menetelmä käsittää, että määritetään biologisesta näytteestä, joka saadaan kohteelta, seuraavien biomarkkerien kvantitatiivinen arvo:

30 - glykoproteiiniasetyyliit,
 - histidiini,
 35 - vähintään yksi rasvahappomäärä (t)
 seuraavista: dokosaheksaeenihapon suhde
 kokonaisrasvahappoihin, linolihapon suhde

kokonaisrasvahappoihin, kertatyydyttymättömien
rasvahappojen ja/tai oleiinihapon suhde
kokonaisrasvahappoihin, omega-3-rasvahappojen suhde
kokonaisrasvahappoihin, omega-6-rasvahappojen suhde
5 kokonaisrasvahappoihin, tyydyttyneiden rasvahappojen
suhde kokonaisrasvahappoihin,
rasvahappotyydyttymättömyysaste; ja
verrataan biomarkkerien kvantitatiivista arvoa
tai kvantitatiivisia arvoja kontrollinäytteeseen tai
10 kontrolliarvoon tai kontrolliarvoihin;
jossa biomarkkerien kvantitatiivisen arvon tai
kvantitatiivisten arvojen nousu tai lasku verrattuna
kontrollinäytteeseen tai kontrolliarvoon
osoittaa/osoittavat kohteen kohonnutta riskiä kuolla
15 rintasyöpään.

5. Jonkin patenttivaatimuksista 1 - 4 mukainen
menetelmä, jossa rintasyöpä käsittää rinnan
pahanlaatuisen kasvaimen tai on rinnan pahanlaatuinen
kasvain (ICD-10-diagnoosikoodi C50).

20 6. Jonkin patenttivaatimuksista 1 - 5 mukainen
menetelmä, jossa vähintään yhden biomarkkerin
kvantitatiivinen arvo mitataan ydinmagneettisella
resonanssispektroskopiolla.

7. Jonkin patenttivaatimuksista 1 - 6 mukainen
25 menetelmä, jossa menetelmä käsittää edelleen, että
määritetään, onko kohteella riski kuolla rintasyöpään,
käyttämällä riskipistemäärää, riskisuhdetta,
kerroinsuhdetta (odds ratio) ja/tai ennustettua
absoluuttista riskiä tai suhteellista riskiä, joka
30 lasketaan vähintään yhden biomarkkerin tai useiden
biomarkkerien kvantitatiivisen arvon tai
kvantitatiivisten arvojen perusteella.

8. Jonkin patenttivaatimuksista 1 - 7 mukainen
menetelmä, jossa kontrollinäyte tai kontrolliarvo
35 edustaa naisia, joilla ei ole tiedossa olevaa
rintasyöpädiagnoosia, tai naisia, joilla on jo
diagnosoitu rintasyöpä.

9. Jonkin patenttivaatimuksista 1 - 8 mukainen menetelmä, jossa kohde on nainen yleisessä väestöseulontatutkimuksessa, kuten mammografiassa.

5 10. Jonkin patenttivaatimuksista 1 - 9 mukainen menetelmä, jossa menetelmä käsittää, että määritetään biologisesta näytteestä, joka saadaan kohteelta, seuraavien biomarkkerien kvantitatiivinen arvo:

- 10 - glykoproteiiniasetyyliit,
 - dokosaheksaeenihapon suhde kokonaisrasvahappoihin,
 - linolihapon suhde kokonaisrasvahappoihin,
 - kertatyydyttymättömien rasvahappojen ja/tai oleiinihapon suhde kokonaisrasvahappoihin,
 - 15 - omega-3-rasvahappojen suhde kokonaisrasvahappoihin,
 - omega-6-rasvahappojen suhde kokonaisrasvahappoihin,
 - tyydyttyneiden rasvahappojen suhde kokonaisrasvahappoihin,
 - 20 - rasvahappotyydyttymättömyysaste,
 - histidiini; ja
- verrataan biomarkkerien kvantitatiivista arvoa tai kvantitatiivisia arvoja kontrollinäytteeseen tai kontrolliarvoon tai kontrolliarvoihin;
- 25 jossa biomarkkerien kvantitatiivisen arvon tai kvantitatiivisten arvojen nousu tai lasku verrattuna kontrollinäytteeseen tai kontrolliarvoon osoittaa/osoittavat kohteen kohonnutta riskiä kuolla
- 30 rintasyöpään.

C50: Malignant neoplasm of breast

Mortality

n = 63 859
311 events

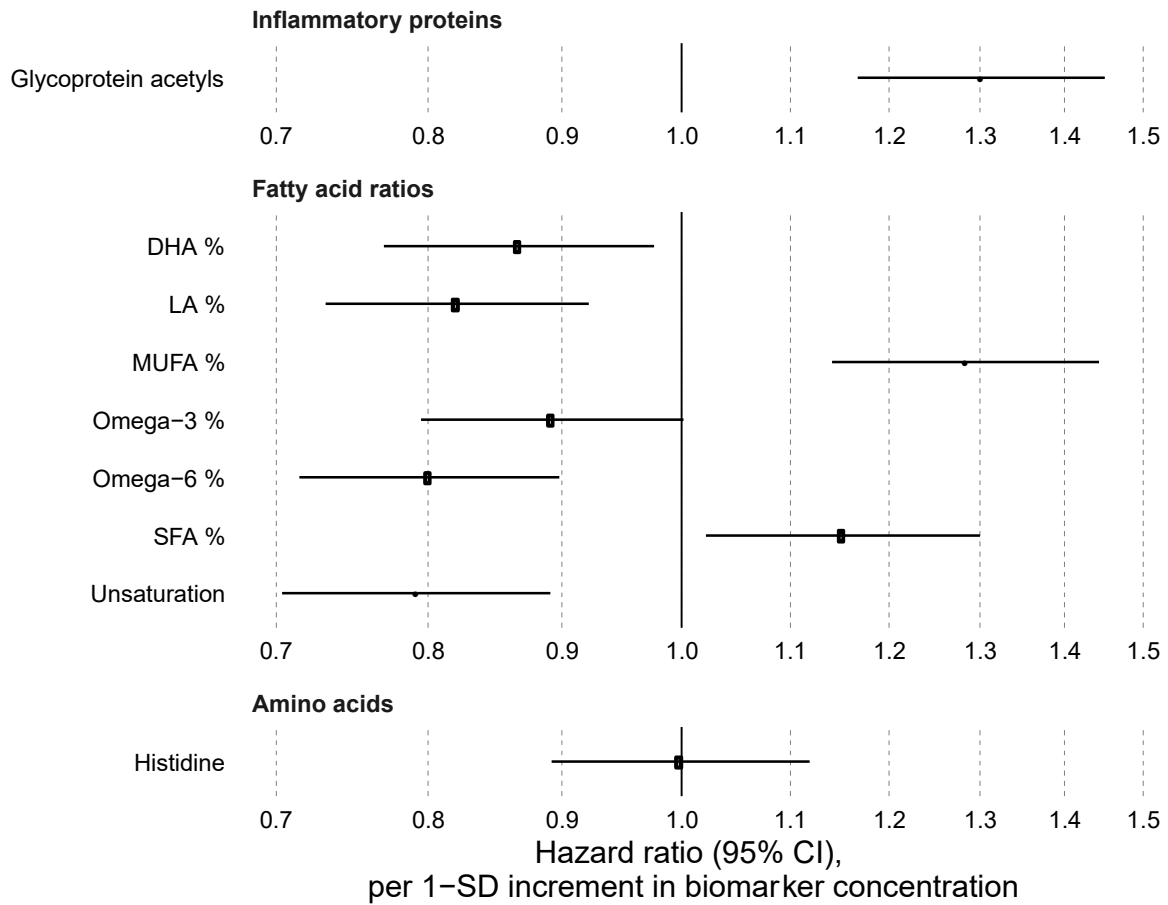


Figure 1

.....

C61: Malignant neoplasm of prostate

Mortality

n = 54 042
256 events

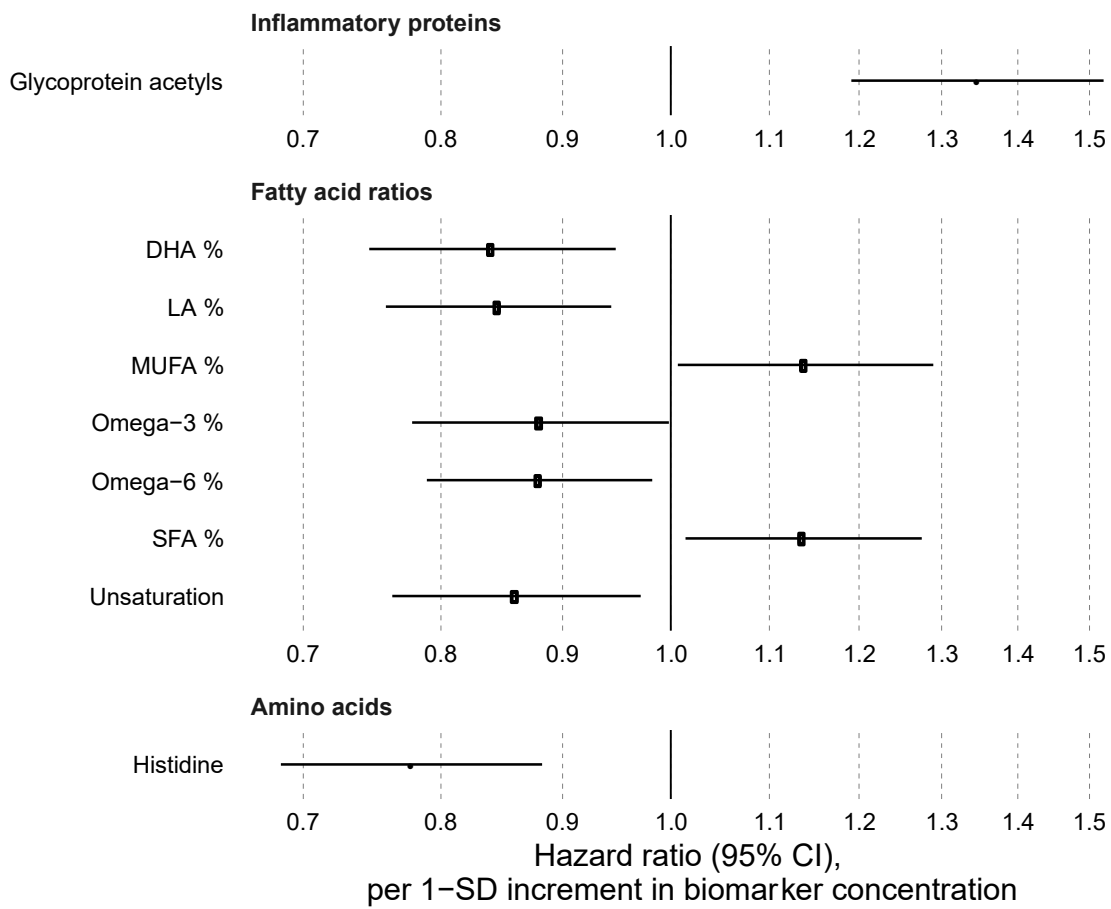


Figure 2

C50: Malignant neoplasm of breast

Mortality among prevalent cases

n = 1536 (prevalent C50)
130 events (death from C50)

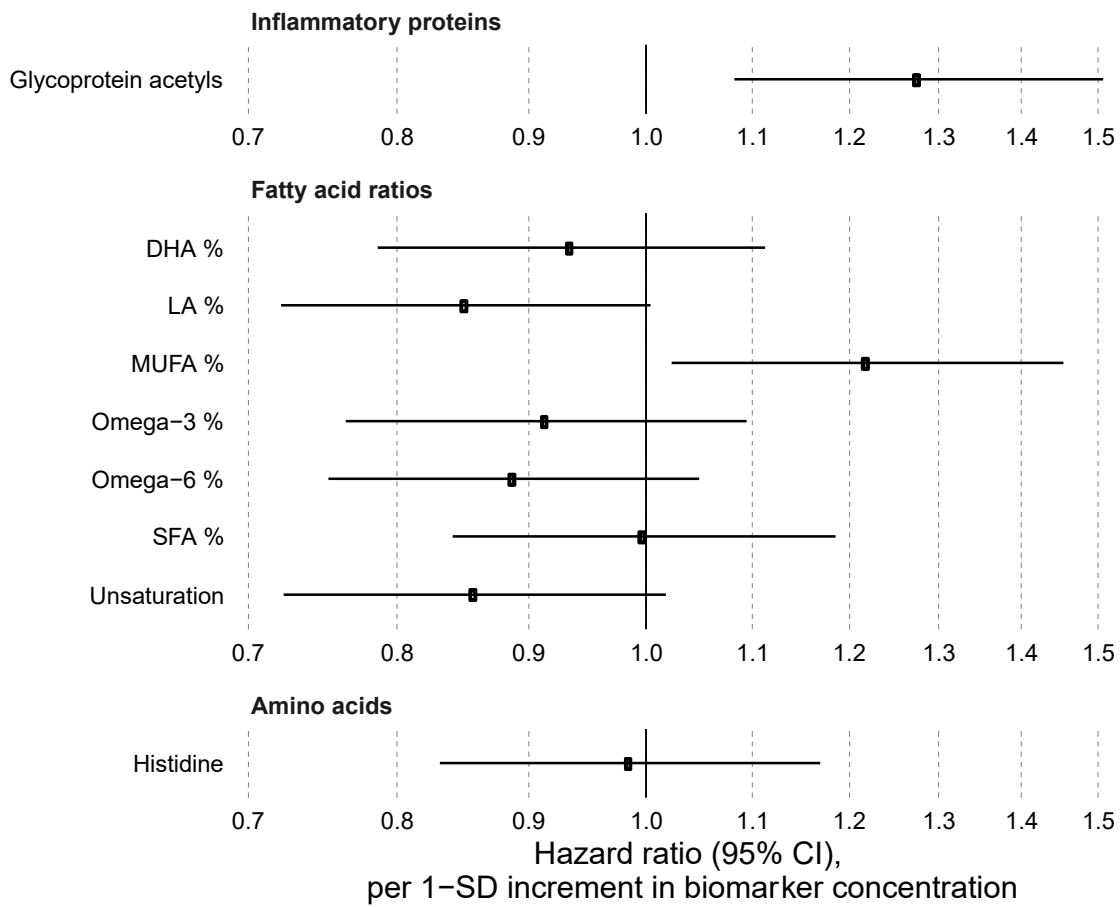


Figure 3

C61: Malignant neoplasm of prostate

Mortality among prevalent cases

n = 584 (prevalent C61)
55 events (death from C61)

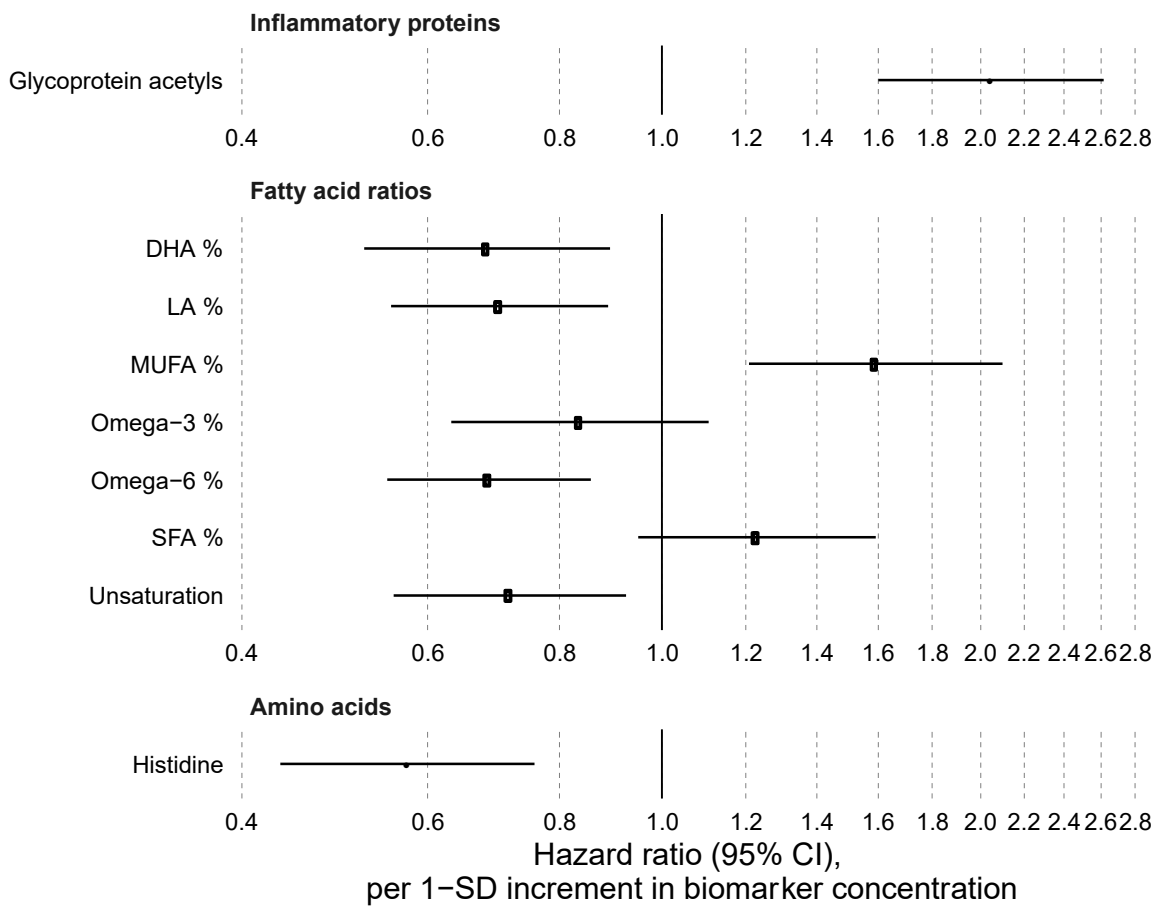


Figure 4

C50: Malignant neoplasm of breast (311 e events)

Mortality risk

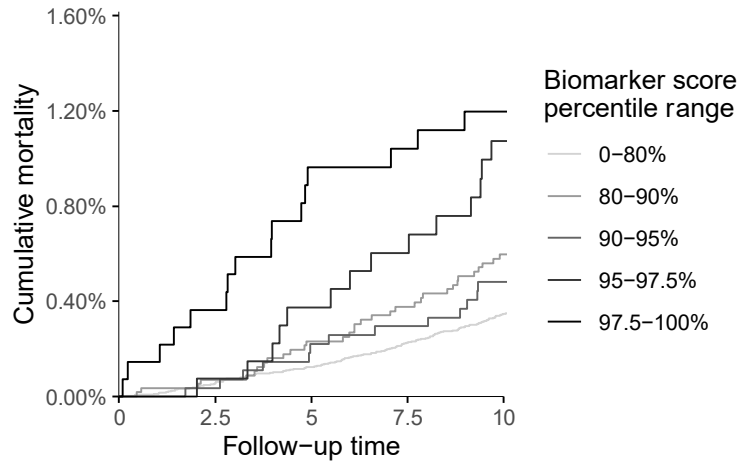


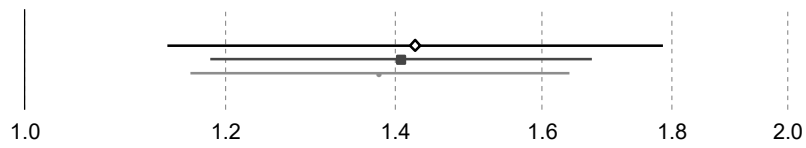
Figure 5a

C50: Malignant neoplasm of breast (311 e events)

Mortality

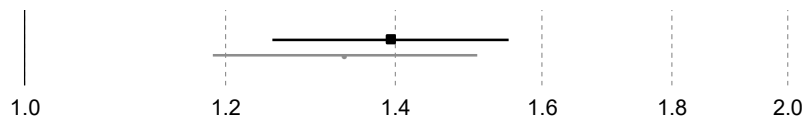
Age at blood sampling

- 61-70
- 53-61
- ◆ 39-53



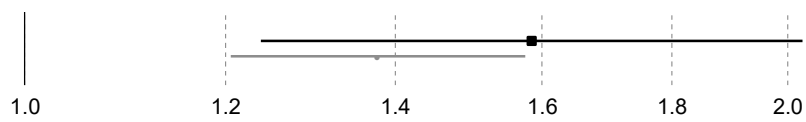
Additional adjustments

- Assessment center, BMI and smoking status
- Assessment center



Short vs. long term risk

- 3-10 years from blood sampling
- Within 3 years from blood sampling



Hazard ratio (95% CI),
per 1-SD increment in multibiomarker score

Figure 5b

C61: Malignant neoplasm of prostate (254 events)

Mortality risk

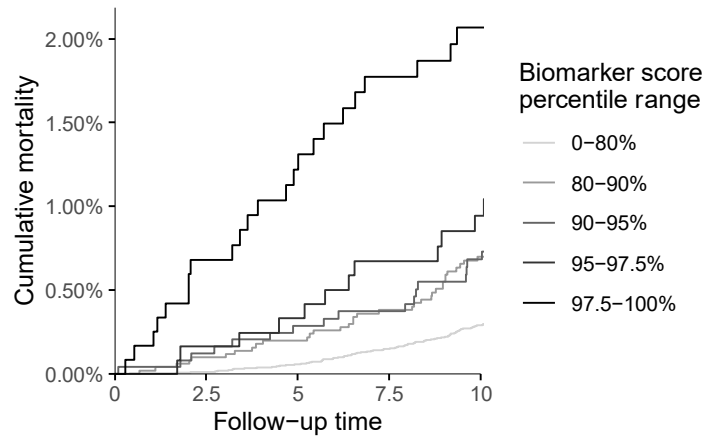


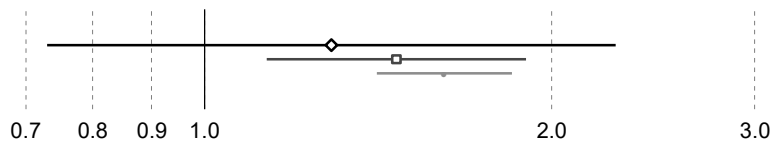
Figure 6a

C61: Malignant neoplasm of prostate (254 events)

Mortality

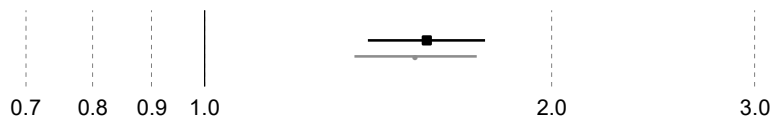
Age at blood sampling

- 61-71
- 53-61
- ◆ 39-53



Additional adjustments

- Assessment center, BMI and smoking status
- Assessment center



Short vs. long term risk

- 3-10 years from blood sampling
- Within 3 years from blood sampling



Hazard ratio (95% CI),
per 1-SD increment in multibiomarker score

Figure 6b