

Lung Cancer Diagnosis

Technical Field

The present invention relates to methods for diagnosis of lung cancer, methods for detecting lung cancer, methods for detecting antibodies, further methods and detection kits and diagnostic test kits.

Prior Art and the Problem Underlying the Invention

Lung cancer is the leading cause of cancer death worldwide. Treatment methods other than surgery are not very efficient and lead to resistance. Thus, insights into the etiology of lung cancer and its progression are urgently needed. Colorectal cancer is another leading cause of cancer-related death and the fourth most common cancer worldwide. The survival and prognosis of colorectal cancer patients depends on the stage of the tumor at the time of diagnosis. Early stages of colorectal cancer can be curable. Unfortunately, over 57% have regional or distant spread of the disease at the time of diagnosis. Despite significant investment and advances in the management of cancer, the five-year survival is only 15% for advanced stage colorectal cancer patients.

Recently, many groups have addressed the mechanisms that drive lung cancer by comparing protein, RNA, and microRNA in tumors with healthy tissue. Besides *TP53*, the most frequently deleted or mutated gene in lung cancer, components of the p53-ARF pathway are also consistently deleted, mutated, or epigenetically modified. As to the colorectal cancer, the challenges are to understand the molecular basis, and to determine factors that initiate the development, and drive the progression. The molecular events involved in colorectal cancer onset and metastatic progression have only been partially clarified. Recent studies have revealed the potential use of molecular and biochemical markers in colorectal cancer to predict outcome and response to chemotherapy, like MLH1, MSH2, β -Catenin, and p53.

Molecular profiles are emerging as predictive and prognostic parameters in non-small-cell lung cancer (NSCLC), including genes involved in DNA damage repair, such as *ERCC1*, *RRM1*, and *BRCA1*. The up-regulated expression of the breast cancer predisposition gene, *BRCA1* was proposed as prognostic and predictive marker for response to treatment in NSCLC. Concerning colorectal cancer, the studies of the *BRCA1* are mainly limited in

colorectal cancer risk and BRCA1 mutations. Several studies attempted to correlate BRCA1 mutations and colorectal cancer risk, but without any clear conclusions. Based on the current limited available evidence, BRCA mutation carriers should be regarded as at high risk for colorectal cancer. However the specific role of BRCA1 expression in colorectal cancer is unclear.

BRCA1 is expressed in many proliferating tissues and acts as a tumour suppressor in DNA repair pathways and cell cycle control. BRCA1 protein stability and function depend on its interaction with BARD1 (BRCA1 associated RING domain protein 1). The BRCA1-BARD1 heterodimer has E3 ubiquitin ligase activity, thus controlling the stability of key target proteins through ubiquitination. BARD1 is also involved in p53-dependent apoptosis, which is deficient in most lung cancers. BARD1 stabilizes p53 and promotes its phosphorylation, and expression of BARD1 is required for proper p53 functioning in signalling towards apoptosis. Thus, BARD1 plays a dual role in tumour suppression, as a binding partner of both BRCA1 and p53. Several studies have shown that BARD1 is upregulated during mitosis, transcriptionally by E2F and posttranslationally by phosphorylation, and importantly, that it is essential for mitosis. According to other studies, both BRCA1 and BARD1 were shown to interact with hMSH2, a gene commonly associated with hereditary nonpolyposis colorectal cancer (HNPCC) and mutations of hMSH2 appear to account for approximately 30-40% of HNPCC. Defects in the BRCA1-hMSH2 signalling process lead to increased susceptibility to tumorigenesis.

WO 98/12327 (Board of Regents, the University of Texas System) discloses several genes, identified in screening assays based upon binding to the breast cancer protein, BRCA1. One of these genes is termed BARD1, a RING protein that interacts with BRCA1 and is envisioned for use in various cancer-related diagnostic and therapeutic methods, particularly those connected with breast, ovarian and uterine cancer.

WO 2008/119802 (Université de Genève) discloses that in gynecological cancers, deletion-bearing isoforms of BARD1 are overexpressed and aberrantly localized to the cytoplasm, and their expression correlated with poor prognosis in breast and ovarian cancer. Structural analysis of these isoforms showed that they lacked the regions that interact with BRCA1 or induce apoptosis. These isoforms are specific to gynecological cancers and are termed as isoforms α , β , η , γ , ϵ , ϕ , δ and Ω .

WO 2012/023112 discloses novel isoforms of BARD1, which are specifically occurring in lung and colon cancers. Methods for detecting these isoforms are disclosed.

Due to the severity of lung cancers, there is an urgent need of providing effective methods for diagnosing lung cancers in a subject. Due to incurability of cancer in subjects being in an advanced state, there is an urgent need for providing methods that allow for early diagnosis of lung cancer. Early diagnosis significantly improves prognosis of subjects suffering from lung cancer.

It is an objective of the invention to provide methods of diagnosis that are non-invasive or that require minimal invasive procedures. A test for diagnosis that can be performed on the basis of a blood sample would be advantageous.

Another objective of the present invention is to provide a rapid, reliable, sensitive and specific test for diagnosis of lung cancer.

The present invention addresses the problems depicted above.

Summary of Invention

The present invention provides methods, kits and assays suitable in diagnosis of lung cancer.

In an aspect, the invention provides a method for detecting and/or measuring levels of antibodies in a blood and/or serum sample, wherein levels of antibodies that specifically bind to different peptides are measured, wherein said peptides comprise a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51).

In an aspect, the invention provides a method for detecting and/or measuring levels of circulating antibodies of a mammalian subject, wherein antibodies that are specific to a number of different peptides are detected, wherein said peptides are related to BARD1 and/or comprise a fragment of an amino acid sequence in BARD1 or any one of its isoforms.

The amino acid sequence of human BARD1 (full length) is provided in the enclosed sequence

listing under SEQ ID NO: 42. Isoforms of BARD1 encompass isoform α (alpha) (SEQ ID NO: 43); isoform π (pi) (SEQ ID NO: 44); isoform β (beta) (SEQ ID NO: 45); isoform κ (SEQ ID NO: 46); isoform γ (SEQ ID NO: 47); isoform λ (SEQ ID NO: 48); isoform ϕ (phi) (SEQ ID NO: 49); isoform ϵ (epsilon) (SEQ ID NO: 50); and isoform η (eta) (SEQ ID NO: 51).

In an aspect, the invention provides a method of diagnosing lung cancer by detecting levels of circulating antibodies in a mammalian subject, wherein antibodies that are specific to a number of different peptides are detected, wherein said peptides are related to BARD1 and/or comprise a fragment of an amino acid sequence in BARD1 or any one of its isoforms.

In an aspect, the present invention provides an *in vitro* and/or *ex vivo* method for measuring and/or detecting levels of circulating antibodies in a mammalian subject, the method comprising the steps of: a) determining, on the basis of a blood or serum sample taken from the subject, a parameter related to the amount of circulating serum antibodies specifically binding to different peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51)

In an aspect, the present invention provides an *in vitro* and/or *ex vivo* method for diagnosing lung cancer in a mammalian subject, the method comprising the steps of: a) determining, on the basis of a blood or serum sample taken from the subject, a parameter related to the amount of circulating serum antibodies specifically binding to different peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51); b) diagnosing, on the basis of the parameters obtained in the previous step, whether said subject suffers or not from lung cancer.

In an aspect, the present invention provides an *in vitro* and/or *ex vivo* method of detecting and/or measuring levels of antibodies associated with the occurrence of lung cancer, the method comprising the step: a) determining, on the basis of said blood or serum sample taken from the subject, parameters related to the amount of antibodies specifically binding different peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51).

In an aspect, the present invention provides a method for assessing a probability and/or a risk that a subject suffers from lung cancer, the method comprising the steps of: a) determining, on the basis of a blood or serum sample taken from the subject, parameters related to the amount of antibodies specifically binding to different peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51); assessing, from the parameters measured in step a) the probability and/or risk that said subject suffers from lung cancer.

In an aspect, the present invention provides the use of a combination of different peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51) in an *in vitro* and/or *ex vivo* in the methods of the invention, for example in a method for lung cancer screening, monitoring, diagnosis, prognosis, prediction, recurrence, and/or in methods for enhancing the clinical efficiency of lung cancer screening, monitoring, diagnosis, prognosis, prediction and recurrence.

In an aspect, the present invention provides a method for detecting and/or determining levels of antibodies present in a blood and/or serum sample, the method comprising the steps of: Providing a blood or serum sample from a mammalian subject; Bringing the said blood or serum sample into contact with one or more surfaces and/or solid matrices where different peptides are bound to, said peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51), wherein the contacting is under conditions sufficient for binding an antibody present in the said sample to the different peptides through antigen-antibody interactions; Removing the blood or serum sample for removing from the one or more surfaces or solid matrices any unbound antibody; Determining levels of an antigen-antibody complex bound to the said surfaces and/or matrices.

In an aspect, the present invention provides a method for diagnosing lung cancer in a mammalian subject, the method comprising the steps of: Providing a blood or serum sample from a mammalian subject; Bringing the said blood or serum sample into contact with one and/or more surfaces or solid matrices where different peptides are bound to, said peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51), wherein the contacting is under conditions sufficient for binding an antibody present in the said sample to the different peptides through

antigen-antibody interactions; Removing the blood or serum sample for removing any unbound antibody from the surfaces and/or matrices; Determining levels of an antigen-antibody complex bound to the said surfaces and/or matrices; and diagnosing, on the basis of the levels obtained in the previous step, whether said subject suffers or not from lung cancer.

5

In an aspect, the invention provides an *in vitro* and/or *ex vivo* method for diagnosing lung cancer in a female human subject. In a preferred embodiment, the method comprises the steps of: Determining, on the basis of a blood or serum sample taken from the female subject, parameters related to the amount of antibodies specifically binding to at least 19 different peptides, wherein said peptides are peptides comprising the amino acid sequences SEQ ID NO: 1 to SEQ ID NO: 6; SEQ ID NO: 11 to SEQ ID NO: 14; SEQ ID NO: 18 to SEQ ID NO: 21; SEQ ID NO: 26 to SEQ ID NO: 30, respectively; and, Diagnosing, on the basis of the parameters obtained in the previous step, whether said female subject suffers or not from lung cancer.

10

15

In an aspect, the invention provides an *in vitro* and/or *ex vivo* method for diagnosing lung cancer in a male human subject. In a preferred embodiment, this method comprises the steps of: Determining, on the basis of a blood or serum sample taken from the male subject, parameters related to the amount of antibodies specifically binding to at least 22 different peptides, wherein said peptides are peptides comprising the amino acid sequences SEQ ID NO: 1 to SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 11 to SEQ ID NO: 13; SEQ ID NO: 15; SEQ ID NO: 18 to SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 26; SEQ ID NO: 27; and SEQ ID NO: 31 to SEQ ID NO: 34, respectively; and Diagnosing, on the basis of the parameters obtained in the previous step, whether said male subject suffers or not from lung cancer.

20

25

In an aspect, the present invention provides an *in vitro* and/or *ex vivo* method for diagnosing lung cancer in a subject, the method comprising the steps of:

- a) - exposing a combination of different peptides to a blood or serum sample of the subject, wherein said peptides are selected from peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) or of an isoform thereof (SEQ ID NO: 43-51) and wherein said peptides have, independently a length of 4 to 300 amino acids;
- determining a parameter related to the amount of antibodies in said sample binding to each of said peptides;

30

b) determining, from the parameter obtained in the previous step, whether said subject is diagnosed positive or negative for lung cancer.

In some aspects, the present invention provides kits for conducting the methods of the invention. Preferably, the kit comprises at least 4 different peptides.

In an aspect, the invention provides a diagnostic test kit for diagnosis of lung cancer, wherein said diagnostic test kit comprises a combination of at least 11 different peptides, each peptide comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51). In an embodiment, said different peptides are selected from the group of peptides comprising an amino acid sequence according to any one of SEQ ID NOs: 1 to 41.

Further aspects and preferred embodiments of the invention are defined herein below and in the appended claims. Further features and advantages of the invention will become apparent to the skilled person from the description of the preferred embodiments given below.

Brief Description of the Drawings

Figure 1 shows the deviance of a fitted model (a logistic regression model here). The model shown is a measure for the quality (in means of discriminative power) of diagnostic kits for lung cancer and is used to compare different models in accordance with embodiments of the invention.

Figures 2 A to C show receiver operating characteristic (ROC) curves and area under the curve (AUC) values (higher is better, maximum = 1) in accordance with embodiments of models of diagnostic tests in accordance with the invention. **(A)** Represents an optimal 25-peptide model (min), AUC=0.966, excellent discriminatory ability; **(B)** 17-peptide model (1se), AUC=0.927, good discriminatory ability; **(C)** 10-peptide model (AUC<0.828); poor discriminatory ability.

Figures 3 A to C illustrate the performance of support vector machine (SVM) classifiers based on 40, 30 and 10 peptides (variables), respectively. Each Figure A to C shows, from left to right, representative ROC curve; box plot showing relative signal values distribution for

cancer and control samples; box plot showing the AUC value distribution in 100 modeling series (mean is ~ 0.83).

Detailed Description of the Preferred Embodiments

5

In some embodiments, the present invention relates to methods and uses of diagnosis, methods of detecting blood and/or serum antibodies of a mammalian subject, methods for detecting and/or measuring levels of antibodies associated with the occurrence of lung cancer, methods for assessing a probability and/or a risk that a subject suffers from lung cancer. The invention also relates to uses and kits.

10

In some embodiments, the methods of the invention comprise the step of detecting and/or measuring levels of antibodies in a blood and/or serum sample. In some embodiments, the methods of the invention comprise the step of detecting and/or measuring levels of circulating antibodies of a mammalian subject. The antibodies are preferably autoimmune antibodies.

15

In an embodiment, the methods of the invention are *ex vivo* and/or *in vitro* methods. The methods of the invention are preferably not conducted directly on the human or animal body. In some embodiments, the methods of the invention comprise a step of providing a blood sample. Preferably, a blood sample is taken, for example a few drops of blood are withdrawn, for example by capillary blood sampling using, for example, a blood lancet. The blood sample may also be provided by venous blood sampling (venopuncture). Preferably, the methods of the invention are conducted on the basis of a sample taken previously from a subject by routine techniques. In some embodiments, the step of blood sampling is not part of the method of the invention.

25

Accordingly, in some embodiments, the methods of the invention comprise a step of providing a blood or serum sample. Preferably, the blood and/or serum sample is taken from a mammalian subject.

30

For the purpose of the present invention, the subject is preferably a mammalian animal or a human. Most preferably, the subject is a human. According to an embodiment, the subject is a human having a history of smoking. Accordingly, the method of diagnosis is preferably specifically designed for humans that are or have been smokers. In a preferred embodiment,

the method of diagnosis is preferably specifically designed for humans that have a long history of heavy smoking.

Levels of antibodies in a blood sample may be measured quantitatively and/or semi-quantitatively, for example. In some embodiments, the method of the invention comprises the step of determining a parameter related to the amount of antibodies in the sample and/or circulating antibodies. Preferably, the parameter is a numerical value, related to the amount of antibodies present in the sample. The numerical value is preferably produced in an automated process, using suitable reading and data processing equipment, as will be exemplified elsewhere in this specification.

In an embodiment, the methods of the invention comprise the step: a) determining, on the basis of a blood or serum sample taken from the subject, a parameter related to the amount of circulating serum antibodies specifically binding to different peptides comprising a stretch of the amino acid sequence of human BARD1 (SEQ ID NO: 42) and/or of an isoform thereof (SEQ ID NO: 43-51). The parameter related to the amount of serum antibodies may represent and/or correspond to the level of antibodies in said sample.

The peptides of the invention are preferably isolated and/or purified peptides. According to a preferred embodiment, the peptides are recombinant peptides. According to an embodiment, the peptides are synthetic peptides. Synthetic chemistry methods, such as solid-phase peptide synthesis, can be used to synthesize the polypeptides according to the invention. Purification of those peptides may be carried out by means of any technique known in the art of protein/peptide purification. Exemplary techniques include ion-exchange chromatography, hydrophobic interaction chromatography, and immune-affinity methods.

In an embodiment of the invention, the step of determining levels of antibodies in a blood sample and/or said step a) comprises the steps of:

- 1) Bringing the said blood or serum sample into contact with one or more surfaces and/or solid matrices where said different peptides are bound to, wherein the contacting is under conditions sufficient for binding an antibody present in the said sample to the different peptides through antigen-antibody interactions;
- 2) Removing the blood or serum sample for removing from the one or more surfaces or solid matrices any unbound antibody from the surfaces and/or matrices;

3) Determining said parameter related to the amount of circulating serum antibodies by measuring levels of an antigen-antibody complex bound to the said surfaces and/or matrices.

The expressions "solid matrix" and "solid matrices" include any solid phase support suitable for carrying out an immunoassay or a method according to the invention. It includes beads, microparticles, nanoparticles, tubes, fabrics or plates, films, slides, wells, formed from or coated with glass, polystyrene, polypropylene, nitrocellulose, quartz, ceramic, dextran or other materials. For example, the solid matrix is in a form of microtiter wells, such as a 96- or a 312-well microtiter plate. In some embodiments, the "solid matrices" and/or surface refers to a carbon material or other materials suitable as electrode materials in electrochemiluminescence assays, for example.

The fixation and/or binding of the peptides according to the invention to the solid matrix, for example in the methods and kits of the invention, may be carried out by adsorption or chemical coupling to a solid phase support. Any means known in the art for immobilizing a protein or peptide to a solid support can be used. The peptides according to the invention can be either covalently or non-covalently bound to the solid matrix by techniques such as covalent bonding via an amide or ester linkage or adsorption. Peptides can be bound using binding pairs such as biotin and avidin or antibody and antigen. After the peptides are affixed to the solid matrix, the solid matrix can be incubated with a blocking solution (containing a blocking protein such as bovine serum albumin) to reduce non-specific adsorption of antibodies in a test sample to the support surface. According to one embodiment, the peptides according to the invention can be synthesized directly on the solid matrix of the kit of the invention, for example.

For the purpose of this specification, the expression "antigen-antibody complex" refers to the complex formed by the binding of the antibodies in the sample to the peptides of the invention, preferably provided on a surface and/or solid support. Therefore, "antigen-antibody complex" may also be referred to as "peptide-antibody complex" for the purpose of this specification. In said step 3), the antibodies retained on the surfaces and/or matrices are preferably those that form the antigen-antibody complex, wherein said peptides function as antigens for said antibodies.

Step 2) of removing said blood and/ or serum sample is preferably conducted by washing

and/or rinsing the one or more surfaces and/or solid matrices, for example with a suitable washing buffer, such as PBS. Preferably, this step does not disrupt the antigen-antibody that may have occurred in step 1).

- 5 In an embodiment, step 3) comprises the step of adding a marker molecule to the one or more surfaces or solid matrices after step 1) and/or after step 2), said marker interacting with and/or binding to the antibodies of the blood or serum sample which form the antigen-antibody complex bound to the said surfaces and/or matrices.
- 10 The marker molecule is preferably an entity that produces or can be induced to produce a signal that can be read in a signal reading step. Generally, the marker molecule associates or is associated in some way with the antibodies, in particular with the antigen-antibody complex. Therefore, the presence of the marker generally indicates the presence of the antigen-antibody complex, and, more specifically, the presence of the antibody that is specific
- 15 to the peptide.

The marker may comprise an antibody or antibody fragment covalently connected to a signal-producing compound or molecule, such as a dye, etc. Alternatively, the marker may be a molecule that can be directly conjugated to the antibodies bound in the form of antigen-antibody complexes on the surfaces and/or solid matrices.

20

There are many types of markers available to the skilled person, such as markers producing light signals (dye), magnetic signals (e.g. magnetic beads) and radioactive signals (e.g. radioactive label for radioimmunoassay), for example.

25

In an embodiment, the method of the invention (for example, step a)) comprises the step of producing and/or measuring a signal that is related to the amount of antibodies in said sample binding to each of respective peptide. In an embodiment, the method comprises a step of measuring a signal that is dependent on the presence and/or quantity of an antigen-antibody complex bound to the said surfaces and/or matrices.

30

In preferred embodiments, the marker molecules can be induced to produce a signal that can be read with suitable reading equipment. Accordingly, in some embodiments, the method of the invention comprises a step of generating and measuring a signal that is dependent on

and/or affected by the presence and/or quantity of an antigen-antibody complex bound to the said surfaces and/or matrices, wherein a quantity of said complex affects the nature, intensity and/or strength of said signal.

- 5 Depending on the marker molecule used, the skilled person will choose an appropriate method for generating the signal, such as, for example, the substrate for an enzyme in case of an ELISA, or the application of electricity in case of electrochemiluminescence, for example.

10 Signal reading is preferably conducted in an automated process, using suitable reading equipment. Preferably, a plate reader is used for measuring the signal. For example, if a light signal is measured, the reading equipment may comprise a light sensor, for example a photomultiplier tube and/or a camera. For example, a charge-coupled device (CCD) camera may be used.

- 15 The reading equipment generally directly calculates or determines, from the measurement of the signal, the numerical parameter that is related to the amount of antibodies in the blood sample. In case of electrochemiluminescence detection (Meso Scale Discovery, USA), SECTOR Imager 6000 and 2400 are commercially available imaging detection systems. Such apparatuses are run with suitable software that directly calculates numerical values from the
20 signal read on the assay plates. In case of an ELISA, for example, conventional ELISA plate readers may be used for determining the numerical parameter that is related to the amount of antibodies.

25 In general terms, methods for conducting step a) and/or for measuring levels of antibodies in a blood sample encompass optical detection methods (e.g. ELISA), mass variation detection (e.g. surface Plasmon resonance, mass spectrometry), and electrical detection (e.g. impedance, spectroscopy, electrochemical) techniques.

30 In an embodiment, levels of antibodies in a blood sample are determined immunochemically, for example by radioimmunoassay, immunofluorescence assay or by an enzyme-linked immunosorbent assay, and immunoassays based on antibodies to protein, for example.

As the skilled person will understand, the use of the word "peptide" in the singular encompasses that a plurality of the same and/or identical peptides are referred to, for example

a plurality of identical peptides (in terms of sequence and/or structure) fixed on one spot of a microtiter plate.

5 The expression "different peptides" generally refers to peptides having different structures and/or amino acid sequences. Generally, the expression "different peptides" more specifically refers to a several pluralities or groups of peptides (a plurality of a plurality), wherein each plurality or group is characterized by the identity of the amino acid sequence and/or structure of the peptides within that plurality and/or group. On the other hand, "different peptides" are contained in different pluralities and/or groups, which are different in that the peptides they
10 contain have a different structure and/or a different sequence. Accordingly, the expression "different peptides" is not intended to mean "different individual peptide molecules having the same amino acid sequence" for the purpose of this specification.

15 In accordance with an embodiment, a parameter related to the amount of antibodies is determined separately for each peptide, so as to obtain a plurality of said parameters in which each individual parameter is related to the amount of antibodies binding to one specific peptide of said different peptides.

20 In an embodiment, said different peptides are a combination of different peptides.

In an embodiment of the methods of the invention, the amount of antibodies is determined separately for each of said different peptides. In particular, for each of the peptides that is part of the method, kit and/or assay of the invention, a value and/or parameter is produced that is related to the amount of antibodies present in the sample. Preferably, for each of the different
25 peptides, the level of antibodies specifically binding to the peptide is determined.

The method of the invention encompasses the use of a plurality of different peptides and measuring levels of antibodies binding to each of the different peptides.

30 In an embodiment, said peptides are provided in one or more wells of a microtiter plate. Preferably, in the methods of the invention, the sample of said subject is added to said one or more wells. In an embodiment, each well comprises several areas and wherein each area comprises a plurality of a specific, defined peptide selected of said combination of different peptides.

In an embodiment, all peptides of said different peptides have, independently, a length of 4 to 300, preferably 8 to 250 amino acids. In preferred embodiments, said peptides have a length of 8 to 200, more preferably 8 to 150, even more preferably 8 to 100 and most preferably 8 to 50 amino acids.

5

In some embodiments, the peptides are short fragments of about 6 to 30 amino acid length, for example and in other embodiments the peptides are larger fragments, having 100-300, preferably 130-260 amino acid lengths, for example. In some embodiments, the peptides used in the methods and kits of the invention comprise shorter and larger fragments and possibly
10 fragments having a length lying in between the above ranges, for example from 30-100 amino acids.

In an embodiment, said different peptides consist of 5 to 35, preferably 11 to 30, and most preferably 17 to 28 different peptides. According to a preferred embodiment, said different
15 peptides comprise eleven (11) or more different peptides.

In a preferred embodiment, said different peptides are selected from a group of forty-one peptides, wherein each peptide of said group of forty-one peptides comprises and or consists essentially of, respectively, one of the forty-one amino acid sequences of the group consisting
20 of SEQ ID NO: 1 to 41. Preferably, said different peptides are a combination of different peptides, wherein said combination is formed of 41 different peptides, respectively, wherein each peptide comprises and/or consists essentially of one of the forty-one amino acid sequences of the group consisting of SEQ ID NO: 1 to 41.

25 For the purpose of the present specification, the expression "consists essentially of" refers to a sequence identity of at least 85%, preferably at least 90%, even more preferably at least 95%, and most preferably 97% or more. For example, "consists essentially of" refers to 98%, 99% and most preferably 99.5% of sequence identity. For the purpose of the present specification, sequence identity percentage is determined by using the basic protein blast on the internet
30 (<http://blast.ncbi.nlm.nih.gov>) with preset standard parameters and database selections. This sequence comparison tool is based on algorithms detailed in the two following publications: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402. Stephen

F. Altschul, John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schäffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", FEBS J. 272:5101-5109.

5 Standard parameters include the selection of blastp (protein-protein BLAST, automatic adjustment of parameters to short input sequences; expect threshold 10, word size 3, use of the matrix BLOSUM62; Gap costs: existence: 11, extension 1; conditional compositional score matrix adjustment, no filters and no masking). Sequence identity of a sequence of comparison with respect to an original sequence is reduced when, for example, any one of the
10 compared or the original sequence lacks amino acid residues, has additional amino acid residues and/or has one or more amino acid residue substituted by another residue. Sequences having as little as 80% sequence identity with any sequence as defined herein may still provide functional, that is, are suitable as peptides in the kits and methods of the invention.

15 In an embodiment, the invention concerns the selection of peptides that can be used in the methods of the invention, in particular for diagnosis of lung cancer. In some embodiments, different groups of peptides are established in accordance with the invention, wherein said different peptides may be selected from one or more of these groups of peptides. Preferably, said peptides are grouped in five (5) groups, which are groups A-E.

20 Accordingly, in an embodiment, said different peptides comprise several groups of peptides, a group A of peptides, a group B of peptides, a group C of peptides, a group D of peptides and a group E of peptides.

25 In an embodiment, said group A comprises peptides comprising or consisting substantially of amino acid sequences selected from the group consisting of: SEQ ID NO: 1 (#286); SEQ ID NO: 2 (#720); SEQ ID NO: 3 (#493); SEQ ID NO: 4 (#68/524); SEQ ID NO: 5 (#140), SEQ ID NO: 6 (#139); SEQ ID NO: 7 (#349); SEQ ID NO: 8 (#117), SEQ ID NO: 9 (#5); and SEQ ID NO: 10 (#-4).

30 In an embodiment, said group B comprises one or more peptides comprising or consisting substantially of amino acid sequences selected from the group consisting of: SEQ ID NO: 11 (#16); SEQ ID NO: 12 (#453); and SEQ ID NO: 13 (EX4..2), SEQ ID NO: 14 (BRCT.2.), SEQ ID NO: 15 (#15), SEQ ID NO: 16 (#523); and SEQ ID NO: 17 (#109).

In an embodiment, said group C comprises one or more peptides comprising or consisting substantially of amino acid sequences selected from the group consisting of: SEQ ID NO: 18 (#117/635); SEQ ID NO: 19 (#368); SEQ ID NO: 20 (BRCT.1.), SEQ ID NO: 21 (EX4..1); SEQ ID NO: 22 (#188), SEQ ID NO: 23 (LINK), SEQ ID NO: 24 (#A21/635); and SEQ ID NO: 25 (RING).

In an embodiment, said group D comprises one or more peptides comprising or consisting substantially of amino acid sequences selected from the group consisting of: SEQ ID NO: 26 (Ank); and SEQ ID NO: 27 (#A20/122), SEQ ID NO: 28 (#54); SEQ ID NO: 29 (#48/522) and SEQ ID NO: 30 (#149); SEQ ID NO: 31 (#73); SEQ ID NO: 32 (#A-4); SEQ ID NO: 33 (#3ORF); and SEQ ID NO: 34 (#557).

In an embodiment, said group E comprises one or more peptides comprising or consisting substantially of amino acid sequences selected from the group consisting of: SEQ ID NO: 35 (#319), SEQ ID NO: 36 (#702); SEQ ID NO: 37 (#175); SEQ ID NO: 38 (#84); SEQ ID NO: 39 (#A29) SEQ ID NO: 40 (#542) and SEQ ID NO: 41 (#309).

In an embodiment, said different peptides comprise one or more peptides selected from said group A and one or more peptides selected from said any one of said groups B, C, D, and/or E.

In an embodiment, said different peptides comprise one or more peptides selected from said group A, one or more peptides selected from said group B, and one or more peptides selected from said group C, D, and/or E.

In an embodiment, said different peptides comprise one or more peptides selected from said group A, one or more peptides selected from said group B, one or more peptides selected from said group C, and one or more peptides selected from said group D and/or E.

In an embodiment, said different peptides comprise three or more peptides selected from said group A, and three or more peptides selected from said any one of said groups B, C, D, and/or E.

In an embodiment, said different peptides comprise three or more peptides selected from said

group A, two or more peptides selected from said group B, and two or more peptides selected from said group C, D, and/or E.

5 In an embodiment, said different peptides comprise three or more peptides selected from said group A, two or more peptides selected from said group B, two or more peptides selected from said group C, and two or more peptides selected from said group D and/or E.

10 In an embodiment, said combination of different peptides comprises five or more peptides selected from said group A; two or more peptides selected from group B, and two or more peptides selected from group C.

In an embodiment, said combination of different peptides lacks any peptide from group E.

15 In an embodiment, said different peptides comprise two or more peptides comprising an amino acid sequence selected from SEQ ID NO: 1-5; one or more peptides selected from peptides SEQ ID NO: 11-13; and one or more peptides selected from SEQ ID NO: 18-20.

20 In an embodiment, said different peptides comprise three or more peptides comprising an amino acid sequence selected from SEQ ID NO: 1-5; two or more peptides selected from peptides SEQ ID NO: 11-13; and two or more peptides selected from SEQ ID NO: 18-20.

25 In an embodiment, said different peptides comprise four or more peptides comprising an amino acid sequence selected from SEQ ID NO: 1-5; three or more peptides selected from peptides SEQ ID NO: 11-13; and three or more peptides selected from SEQ ID NO: 18-20.

In an embodiment, said combination of different peptides comprises eleven (11) or more different peptides, wherein said peptides comprise, respectively, the amino acid sequences of SEQ ID NO 1-5; 11-13 and 18-20.

30 In an embodiment, said combination of different peptides comprises at least 10 different peptides, wherein each of said 10 peptides, respectively, comprises or consists essentially of one amino acid sequence selected from the group consisting of: SEQ ID NO: 1 to 10.

In an embodiment, said combination of different peptides comprises at least 17 different

peptides, wherein each of said 17 peptides, respectively, comprises or consists essentially of one amino acid sequence selected from the group consisting of: SEQ ID NO: 1 to SEQ ID NO: 17.

5 In an embodiment, said combination of different peptides comprises at least 25 different peptides, wherein each of said 25 peptides, respectively, comprises or consists essentially of one amino acid sequence selected from the group consisting of: SEQ ID NO: 1 to SEQ ID NO: 25.

10 In an embodiment, the method of the invention is a method for diagnosing lung cancer in a human female subject. In another embodiment, the method of the invention is a method for diagnosing lung cancer in a human male subject.

If the method is specifically directed to a human female subject, the different peptides
15 comprise preferably nineteen (19) different peptides, wherein said different peptides comprise the amino acid sequences SEQ ID NO: 1 to SEQ ID NO: 6; SEQ ID NO: 11 to SEQ ID NO: 14; SEQ ID NO: 18 to SEQ ID NO: 21; SEQ ID NO: 26 to SEQ ID NO: 30, respectively.

If the method is specifically directed to a human male subject, the different peptides comprise
20 preferably twenty-two (22) different peptides, wherein said different peptides comprise the amino acid sequences SEQ ID NO: 1 to SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 11 to SEQ ID NO: 13; SEQ ID NO: 15; SEQ ID NO: 18 to SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 26; SEQ ID NO: 27; and SEQ ID NO: 31 to SEQ ID NO: 34, respectively.

25 In an embodiment, the method of the invention comprises a step of diagnosing whether said subject suffers or not from lung cancer. The diagnosis is preferably made on the basis of the levels of antibodies in said serum and/or blood sample and/or on the basis of the parameter related to the amount of circulating serum antibodies.

30 In an embodiment, the method of the invention comprises the step of calculating a test value for the subject on the basis of said parameter and/or said signal and from a statistically determined coefficient specific to each peptide. The method of the invention may comprise the step of diagnosing, by comparing said test value with a threshold value, whether said

subject suffers or not from lung cancer.

In another embodiment, the method of the invention comprises the step of assessing the probability and/or a risk that a subject suffers from lung cancer by comparing said test value
5 with one or more threshold values.

In another embodiment, the method of the invention comprises the step of enhancing the clinical efficiency of lung cancer screening, monitoring, diagnosis, prognosis, prediction and recurrence by comparing said test value with one or more threshold values. A result or
10 outcome of said screening, monitoring, diagnosis, prognosis, prediction is then dependent on whether the test value is higher or lower than the threshold value. For example, a test value exceeding the threshold value generally indicates and/or enhances the clinical efficiency of diagnosis, prediction and/or prognosis that the subject suffers from lung cancer.

15 The coefficient for each peptide is preferably determined statistically. In an embodiment, the coefficient for each peptide is determined using the "Lasso-model", which is disclosed in: *Ribbing J, Nyberg J, Caster O & Jonsson EN (2007) The lasso—a novel method for predictive covariate model building in nonlinear mixed effects models. J Pharmacokinet Pharmacodyn 34: 485–517*. Specific coefficients for preferred peptides are disclosed in the
20 examples herein below.

The levels of the antibodies and/or the parameter related to the amount of each antibody are preferably transformed into a numerical value. On the basis of the coefficient for each peptide of an assay and the numerical value determined for each peptide separately, the test value
25 specific to a given subject is determined.

The step of diagnosis is done by comparing the test value of the subject with the predetermined threshold value. The threshold value is preferably also determined statistically using, for example, the lasso model. In particular, the threshold value can be determined
30 statistically by using blood and/or serum samples from a number of individuals from which it is known whether or not they suffer from lung cancer. For determining the threshold value, the method used for measuring levels of antibodies in the blood and/or serum sample is preferably taken into account.

In some aspects, the present invention relates to a test kit. The expression "kit" comprises at least one peptide according to the invention or a variant thereof or a combination thereof as described herein to be coupled or already coupled to a solid matrix and optionally instructional material. Preferably, the kit comprises a solid support surface on which the
5 different peptides are deposited and/or fixed. For example, the surface is provided on a plate, film, slides and/or wells. Preferably, the support for the peptides used in the kit is adapted to perform an antibody binding assay. In particular, the support is adapted to bind the peptides, for example covalently, and to be exposed to the blood and/or serum sample. Preferably, the support surface and/or matrix of the kit is preferably suitable to be washed for removing
10 unbound antibody and to be exposed to a marker. The support and/or matrix comprising the peptides is preferably adapted to be read by a suitable apparatus for detecting and measuring levels and/or amounts of peptide-antibody complexes formed by non-covalent interactions.

The peptides of the kit are preferably the same as those used in the methods of the invention.
15 Therefore, the embodiments and preferred embodiments defined elsewhere in this specification in particular with the methods of the invention also apply to the kit of the invention. This applies in particular to the selection of peptides used, but also to general peptide characteristics, such as size and number.

20 In an embodiment, the kit is a kit for carrying out a method according to the invention.

In an embodiment, the test kit of the invention comprises a microtiter plate comprising a plurality of wells, wherein each well comprises a plurality of areas, and wherein a plurality of one specific peptide is provided in a specific area of said well, so that each area is
25 characteristic of a specific peptide. For example, the kit of the invention comprises a multi-spot and/or multi-array plate, such as those that are commercially available from Meso Scale Discovery, Inc., USA. For example, each well of the plate may comprise several spots, wherein each spot which is or can be coated with a peptide.

30 Examples

Example 1: Peptide choice and synthesis

Forty-one different peptides comprising SEQ ID NO: 1 to 41 were defined and selected

following studies on expression of isoforms of BARD1 (SEQ ID NO: 42 – 51). These experiments involved siRNA-based specific repression of one isoform and confirmation of the repression of the potential translation product on Western blots, and overexpression of the isoform and confirmation of the expression of the endogenous protein of the correct size.

5

Studies on isoform γ (SEQ ID NO: 47), for example, confirmed that this isoform expresses a protein encoded in exons 1 to 3 (deletion of exon 4 and exons 5-11 not expressed due to stop codon in exon 5). Based on these studies, forty-one peptides were defined (SEQ ID NO: 1 to 41)

10

The forty-one different peptides of SEQ ID NO: 1 to 41 were synthesized by standard peptide solid phase synthesis procedures known to those skilled in the art. Purity of the peptides was at least 80%. Peptides were dissolved in stock solution in 1 mg/mL in buffer and stored in aliquots of 200 μ L at -20°C. Peptides were stored in buffer carbonate, pH 9.6.

15

Most of the peptides of SEQ ID NO: 1 to 41 are short fragments of 8 to about 30 amino acid length. Larger fragments are SEQ ID NO: 13 (EX4..2), SEQ ID NO: 14 (BRCT.2.), SEQ ID NO: 20 (BRCT.1.), SEQ ID NO: 21 (EX4..1), SEQ ID NO: 23 (LINK), SEQ ID NO: 26 (Ank).

20

Example 2: Preparation of assay using the peptides

25

The 41 peptides of Example 1 were deposited on microtiter wells using the Meso Scale MSD technology platform (MesoScale, MD 20850-3173, Rockville, USA). In particular, each peptide was deposited at a determined amount on a spot of multi spot plates. Each multi spot array comprises a carbon-coated working electrode on which the peptides are deposited. For detection of antibody binding, electric stimulation results in generation of light via a sulfo-tag marker molecule (ruthenium II tris-bipyridine -(4-methylsulfonate) NSH ester. The plating of the peptides was performed by MesoScale. Each well contains 10 different spots, so that 10 different peptides were plated in one well of a 96-well microtiter plate.

30

Example 3: Analysing blood samples of patients and control subjects

Blood samples were collected from lung (n=178), colon (n=80), benign breast (n=9),

malignant breast (n=14), benign ovarian (n=50), malignant ovarian (n=43), and neuroblastoma (n=20) human cancer patients and healthy controls (n=266). Gender and age was known from a majority of the subjects.

- 5 Blood samples were analyzed using the peptide-coated microtiter plates described in Example 2. Using the Meso Scale SECTOR Imager 2400 (SI2400) apparatus and the Discovery Workbench 3.0 Software, numerical values related to the amount of antibodies in the blood sample are produced.
- 10 Data analysis showed that not all patients' sera contained antibodies against the same epitopes of BARD1, but that there was a wide-spread distribution of different combinations of epitopes /peptides that were positive in cancer patients. The most distinctive and least distinctive peptides were determined from these results.
- 15 To confirm that antibodies reacted specifically with peptides presented on BARD1 isoforms, larger fragments were spotted in the same way as the shorter peptide fragments. Reactivity of serum antibodies with larger fragments confirmed the results observed with the peptides.

It was found that a combination of short peptides and larger fragments of BARD isoforms
20 allows detecting more cancer patients.

Example 4: Statistical analysis and model building

- The data obtained in Example 3 were statistically analyzed. The data set consists of 379
25 unique samples for which at least one set of variables has been measured. For some samples, 2 or three separate measurements were made and the mean was taken. A variable is a numerical value obtained from the signal measured when exposing a blood sample to one of 40 peptides. In other words, the variable is a value related to the amount of antibodies in the sample specifically binding to one of the peptides. In total, the data set contains measurements
30 for 40 variables, which are divided into four subsets of 10 variables each (10 different peptides are plated as 10 different spots in one well). Each plate contains measurements for peptides from one such subset, and four plates have been analyzed for each subset. The plates will be referred to by the date when they were analyzed. For some samples, only one subset of variables was measured while for others, there are measurements for several subsets. There

are also some samples for which the same variables have been measured several times, on different plates. Information about gender and age are available for the majority of the studied samples. For the cancer samples more detailed diagnostic information is also provided.

- 5 Below an overview of the analysis outcome focusing on the predictive power of the test system in discriminating between control (healthy) and lung cancer blood serum samples is given. The differences in the methodology and analysis results are discussed.

The “*Lasso*” method was recently used to build models using different peptides selection and then compared the predictive abilities of these models (Fig. 1). When analyzing a small dataset *stepwise covariate modeling procedure (SCM)* may produce a covariate model that suffers from selection bias and poor predictive performance. Compared to the SCM, the *lasso* is superior to SCM in obtaining a predictive covariate model on a small dataset or on small subgroups (Ribbing J, Nyberg J, Caster O & Jonsson EN (2007) *The lasso—a novel method for predictive covariate model building in nonlinear mixed effects models. J Pharmacokinet Pharmacodyn* 34: 485–517). Using cross-validation, the lasso provides a validation of the covariate model and does not require the user to specify a P-value for selection. This method allowed obtaining the optimal combination of oligopeptides and BARD1 fragments for discrimination between cancer and control samples. That finally leads to the selection of the best peptide combination that allow distinguish between cancer and control samples. The figure also demonstrates that increasing the number of peptides over the optimal does not produce stronger prediction power. The modeling with the use of optimal 25 variables (short peptides and larger fragments) performed for the 90 lung cancer samples versus the 94 control samples yielded AUC=0.966.

25

Lasso model building and feature selection was cross-validated using the R-package: glmnet. Three models have been built with 25 (SEQ ID NO: 1-25), 17 (SEQ ID NO: 1-17) or 10 (SEQ ID NO: 1-10) peptides.

- 30 Figure 1 shows the binomial deviance of a fitted model. The number of peptides used in the model is shown on the top. The best cross-validated model is the minimum (first vertical dotted line) of this curve (26 peptides). The second dotted line is the value where the model is not significantly different (inside 1 standard error (1se), shown as error-bars) concerning cross-validation to the models (in this case 11 peptides). Therefore the model with 17 peptides

is statistically meaningful, but not the model with 10 peptides.

The modeling with the use of 25 variables (shorter peptides and larger fragments) (SEQ IF NO: 1-25) performed for 90 lung cancer samples versus 94 control samples yielded AUC=0.966 (Figure 2A). These samples for the modeling were selected from the overall 379 available samples mentioned above. The modeling was also done with the aim to find the minimal set of peptides sufficient to discriminate between cancer and control with the $p > 0.05$. For the AOC curve shown in Figure 2B, the model with 17 peptides (SEQ IF NO: 1-17) was used, and for Figure 3C a model with 10 peptides (SEQ IF NO: 1-10).

Using the Lasso method covariate model building was built with the aim of discriminating between control subjects and cancer patients. The results are shown in Figures 3 A to C using 40, 30 and 10 peptides in Figures A, B and C, respectively in a 100 modeling series (modeling was repeated 100 times). Each Figure A to C shows, from left to right, representative ROC curve; box plot showing relative signal values distribution for cancer and control samples; box plot showing the AUC value distribution in 100 modeling series (mean AUC is ~ 0.83).

Example 5: Peptide coefficients and cutoff value for model with 25 peptides

This example illustrates the calculation of a predictor value is based on coefficients attributed to each peptide as shown in Table 1 below.

Predictor= 1.878155 x log10 (#16 value) -49.108 x log10 (#117 value) -16.4289 x log10 (#286 value) +24.11767 x log10 (#493 value) -3.74674 x log10 (#523 value) +39.89181 x log10 (#720 value) +3.49733 x log10 (#117/635 value) -3.21744 x log10 (#5 value) -13.1504 x log10 (#349 value) -19.78 x log10 (#453 value) +10.3537 x log10 (A21/635 value) +27.37549 x log10 (#68/524 value) -12.5612 x log10 (#15 value) -7.82582 x log10 (#109 value) +15.05732 x log10 (#139 value) +11.3804 x log10 (#140 value) +4.024697 x log10 (#188 value) -8.8073 x log10 (#368 value) +0.246625 x log10 (RING value) -0.23952 x log10 (EX4.1 value) -5.00989 x log10 (EX4.2 value) -0.17352 x log10 (LINK value) +0.726563 x log10 (BRCT1 value) +1.645815 x log10 (BRCT2 value) +3.05931 x log10 (#-4 value) -15.0948

The values, for example "#16 value" correspond to the numerical value of the parameter

determined by the Meso Scale SECTOR Imager. By inserting the numerical value in the formula "Predictor" above, a predictor value is obtained, which is compared to the cutoff and/or threshold value.

5 **Table 1: 25 peptides model**

25 Peptides MODEL		
Peptide identification (SEQ ID NO)	Peptide identification (internal)	Coefficient
	(Intercept)	-15.0948
11	#16	1.878155
8	#117	-49.108
1	#286	-16.4289
3	#493	24.11767
16	#523	-3.74674
2	#720	39.89181
18	#117/635	3.49733
9	#5	-3.21744
7	#349	-13.1504
12	#453	-19.78
24	#A21/635	10.3537
4	#68/524	27.37549
15	#15	-12.5612
17	#109	-7.82582
6	#139	15.05732
5	#140	11.3804
22	#188	4.024697
19	#368	-8.8073
25	RING	0.246625
21	EX4..1.	-0.23952
13	EX4..2.	-5.00989
23	LINK	-0.17352
20	BRCT.1.	0.726563
14	BRCT.2.	1.645815
10	#-4	3.05931

Various cutoffs were estimated and might be used to distinguish between positive or negative. The choice of each cutoffs affect the specificity and/or sensitivity of the model. For a Maximum Specificity and Sensitivity, Cutoffs = 0.08131021.

The cutoff value represents the threshold value for the outcome of the diagnostic test. If the

calculated predictor is >0.08131021 , the subject is diagnosed with cancer (positive test outcome). If the calculated predictor is <0.08131021 , the test result is negative.

Example 6: Model for discriminating lung cancer versus healthy in women

5

Using the Lasso method described in Example 4, a model was developed for diagnosing lung cancer specifically in women. For this model, 19 peptides were found to be required to obtain a model that is statistically significant and results in excellent AUC values. The 19 different peptides are the peptides having the sequences: SEQ ID NO: 1 to SEQ ID NO: 6; SEQ ID
10 NO: 11 to SEQ ID NO: 14; SEQ ID NO: 18 to SEQ ID NO: 21; SEQ ID NO: 26 to SEQ ID NO: 30.

Example 7: Model for discriminating lung cancer versus healthy in men

15 Using the Lasso method described in Example 4, a model was developed for diagnosing lung cancer specifically in men. For this model, 22 peptides were found to be required to obtain a model that is statistically significant and results in excellent AUC values. The 22 different peptides are the peptides having the sequences SEQ ID NO: 1 to SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 11 to SEQ ID NO: 13; SEQ ID NO: 15; SEQ ID NO: 18 to
20 SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 26; SEQ ID NO: 27; and SEQ ID NO: 31 to SEQ ID NO: 34.

Conclusion

25 We have analyzed 178 lung cancer samples and 266 control samples. The dataset consists of 379 unique samples for which at least one set of variables (peptides or polypeptides) has been measured. In total, the data set contains measurements for 40 variables, which are divided into four subsets of 10 variables each.

30 From the raw data of the 40 peptides, an independent analysis using the "Lasso" method was performed to statistically evaluate different peptide selections. It was found that, for example, using 20 selected peptides a predictive performance of $>95\%$ (0.97 ROC) for lung cancer diagnosis is attained.

This is a highly conclusive result which gives all indications of accuracy (>95%) sufficient for predictions to the level of diagnosis of lung cancer in asymptomatic individuals. This applies in particular to individuals with a long history of heavy smoking.

CLAIMS:

1. An *in vitro* and/or *ex vivo* method for diagnosing lung cancer in a mammalian subject, the method comprising the steps of:

a) determining, on the basis of a blood or plasma sample taken from the subject, parameters related to the reactivity of antibodies specifically binding to a combination of at least 11 different peptides comprising a stretch of the amino acid sequence of human BARD1 and/or of an isoform thereof;

b) diagnosing, on the basis of the parameters obtained in the previous step, whether said subject suffers or not from lung cancer;

wherein the combination of at least 11 different peptides is selected from a group of forty-one peptides, wherein each peptide of said group of forty-one peptides comprises and/or consists essentially of, respectively, one of the forty-one amino acid sequences of the group consisting of SEQ ID NO: 1 to 41 wherein step b) comprises the step of calculating a test value for the subject on the basis of said parameters and from a statistically determined coefficient specific to each peptide.

2. Use of a combination of at least 11 different peptides comprising a stretch of the amino acid sequence of human BARD1 and/or of an isoform thereof in an *in vitro* and/or *ex vivo* method for lung cancer screening, monitoring, diagnosis, and prognosis;

wherein the combination of at least 11 different peptides is selected from a group of forty-one peptides, wherein each peptide of said group of forty-one peptides comprises and/or consists essentially of, respectively, one of the forty-one amino acid sequences of the group consisting of SEQ ID NO: 1 to 41.

3. The method of any one of the preceding claim, wherein a parameter related to the reactivity of antibodies is determined separately for each peptide, so as to obtain a plurality of said parameters in which each individual parameter is related to the reactivity of antibodies binding to one specific peptide of said different peptides.

4. The method of claim 1, further comprising the step of diagnosing, by comparing said test value with a threshold value, whether said subject suffers or not from lung cancer.
5. The method of any one of the preceding claims, wherein said peptides have, independently, a length of 4 to 300, or 8 to 250 amino acids.
6. The method of any one of the preceding claims, wherein said different peptides consist of 11 to 35, or 11 to 30, or 17 to 28 different peptides.
7. The method of any one of the preceding claims, wherein said different peptides comprise several groups of peptides, a group A of peptides, a group B of peptides, a group C of peptides, a group D of peptides and a group E of peptides, wherein said group A comprises one or more peptides comprising or consisting of amino acid sequences selected from the group consisting of: SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 6; SEQ ID NO: 7; and SEQ ID NO: 8; SEQ ID NO: 9; and SEQ ID NO: 10, said group B comprises one or more peptides comprising amino acid sequences selected from the group consisting of: SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14 and SEQ ID NO: 15; SEQ ID NO: 16; and SEQ ID NO: 17, said group C comprises one or more peptides comprising amino acid sequences selected from the group consisting of: SEQ ID NO: 18; SEQ ID NO: 19; SEQ ID NO: 20; SEQ ID NO: 21; SEQ ID NO: 22 and SEQ ID NO: 23; SEQ ID NO: 24; and SEQ ID NO: 25, said group D comprises one or more peptides comprising amino acid sequences selected from the group consisting of: SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; SEQ ID NO: 29 and SEQ ID NO: 30; SEQ ID NO: 31; SEQ ID NO: 32; SEQ ID NO: 33 and SEQ ID NO: 34, said group E comprises one or more peptides comprising amino acid 30 sequences selected from the group consisting of: SEQ ID NO: 35; SEQ ID NO: 36; SEQ ID NO: 37; SEQ ID NO: 38; SEQ ID NO: 39; SEQ ID NO: 40 and SEQ ID NO: 41, said combination of peptides comprises one or more peptides selected from said group A and one or more peptides selected from said any one of said groups B, C, D, and/or E.

8. The method of claim 7, wherein said different peptides comprise five or more peptides selected from said group A; two or more peptides selected from group B, and two or more peptides selected from group C.
9. An *in vitro* and/or *ex vivo* method of claim 1 for diagnosing lung cancer in a female human subject, the method comprising the steps of:
- a) determining, on the basis of a blood or plasma sample taken from the female subject, parameters related to the reactivity of antibodies specifically binding to at least 19 different peptides, wherein said peptides are peptides comprising the amino acid sequences SEQ ID NO: 1 to SEQ ID NO: 6; SEQ ID NO: 11 to SEQ ID NO: 14; SEQ ID NO: 18 to SEQ ID NO: 21; SEQ ID NO: 26 to SEQ ID NO: 30, respectively;
 - b) diagnosing, on the basis of the parameters obtained in the previous step, whether said female subject suffers or not from lung cancer wherein step b) comprises the step of calculating a test value for the subject on the basis of said parameters and from a statistically determined coefficient specific to each peptide.
10. An *in vitro* and/or *ex vivo* method of claim 1 for diagnosing lung cancer in a male human subject, the method comprising the steps of:
- a) determining, on the basis of a blood or serum sample taken from the female subject, parameters related to the reactivity of antibodies specifically binding to at least 22 different peptides, wherein said peptides are peptides comprising the amino acid sequences SEQ ID NO: 1 to SEQ ID NO: 5; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 11 to SEQ ID NO: 13; SEQ ID NO: 15; SEQ ID NO: 18 to SEQ ID NO: 20; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 26; SEQ ID NO: 27; and SEQ ID NO: 31 to SEQ ID NO: 34, respectively;
 - b) diagnosing, on the basis of the parameters obtained in the previous step, whether said male subject suffers or not from lung cancer wherein step b) comprises the step of calculating a test value for the subject on the basis of said parameters and from a statistically determined coefficient specific to each peptide.