CLAIMS

1. A method for predicting the likelihood of locoregional recurrence (LRR) and/or distant metastasis in a subject with breast cancer following mastectomy and/or breast conserving surgery, comprising:

    i. measuring the expression level of at least one gene in a sample isolated from the subject; and

    ii. deriving a score based on the measured expression level of the at least one gene using a predictive classification model, wherein the predictive classification model comprises at least one scoring algorithm;

wherein the at least one gene is selected from a group consisting of: TRPV6, DDX39, BUB1B, CCR1, STIL, BLM, C16ORF7, TPX2, PTI1, TCF3, CCNB1, DTX2, PIM1, ENSA, RCHY1, NFATC2IP, OBSL1 and MMP15;

wherein the at least one scoring algorithm is selected from a group consisting of:

    i. Score = TRPV6 + DDX39 + BUB1B + CCR1 + STIL + BLM + C16ORF7 + TPX2 + PTI1 + TCF3 + CCNB1 + DTX2 + PIM1 + ENSA + RCHY1 + NFATC2IP + OBSL1 + MMP15; wherein each gene counts as one point when the hazard ratio is <1, and the total score = 18;

    ii. Score = TRPV6 + DDX39 + BUB1B + CCR1 + STIL + BLM + 2×C16ORF7 + TPX2 + PTI1 + TCF3 + CCNB1 + DTX2 + PIM1 + ENSA + RCHY1 + 2×NFATC2IP + OBSL1 + 2×MMP15; wherein when the hazard ratio is <1, each gene counts as one point if the genes are examined by univariate analysis, and each gene counts as two points if the genes are examined by multivariate analysis, and so the total score = 21;

iii. Score = 3×TRPV6 + 5×DDX39 + 18×BUB1B + 4×CCR1 + 4×STIL + 7×BLM + 7×C16ORF7 + 4×PMI1 + 9×TPX2 + 8×PTI1 + 3×TCF3 + 7×CCNB1 + 2×DTX2 + 2×ENSA + 5×RCHY1 + 6×NFATC2IP + 2×OBSL1 + 6×MMP15; wherein the genes are examined by univariate analysis and the score of each gene is re-scaled according to its weighting, and so the total score = 102;

iv. Score = 2×TRPV6 + 2×DDX39 + 2×BUB1B + CCR1 + STIL + 2×BLM + 5×C16ORF7 + 3×PIM1 + 3×TPX2 + 5×PTI1 + TCF3 + CCNB1 + DTX2 + 2×ENSA + 3×RCHY1 + 4×NFATC2IP + OBSL1 + MMP15; wherein the genes are examined by multivariate analysis and the odds ratio of each gene is re-scaled to a score between 1 and 5, each gene counts as 1 point when the odds ratio is <1, and so the total score = 40; and

v. Score = 4×TRPV6 + 3×DDX39 + 8×BUB1B + CCR1 + STIL + 3×BLM + 11×C16ORF7 + 4×PIM1 + TPX2 + 2×PTI1 + 2×TCF3 + CCNB1 + DTX2 + 2×ENSA + 5×RCHY1 + 4×NFATC2IP + OBSL1 + 2×MMP15; wherein the genes are examined by multivariate analysis and the odds ratio of each gene is re-scaled to a score between 1 and 11, each gene counts as 1 point when the odds ratio is <1, and so the total score = 56; and

wherein the derived score provides an indication of the likelihood of LRR and/or the likelihood of distant metastasis in the subject.

2. The method according to claim 1, wherein the method comprises a step of classifying the subject into a low risk group of LRR and/or distant metastasis when the derived score is less than a first pre-determined reference.

3. The method according to claim 2, wherein the method comprises a step of classifying the subject into a high risk group of LRR and/or distant metastasis when the derived score is equal to or more than the first pre-determined reference.

4. The method according to claim 1, wherein the method comprises a step of classifying the subject into a low risk group of distant metastasis and/or LRR when the derived score is less than a second pre-determined reference.

5. The method according to claim 4, wherein the method comprises a step of classifying the subject into a high risk group of distant metastasis and/or LRR when the derived score is equal to or more than a third pre-determined reference.

6. The method according to claim 4 or 5, wherein the method comprises a step of classifying the subject into an intermediate risk group of distant metastasis and/or LRR when the derived score is between the second pre-determined reference (inclusive) and the third pre-determined reference.

7. The method according to any of the preceding claims, wherein the step of measuring the expression level of the at least one gene comprises hybridizing the at least one gene with at least one gene probe and measuring the expression level of the at least one gene.

8. The method according to claim 7, wherein the at least one gene probe comprises at least one gene selected from a group consisting of: TRPV6, DDX39, BUB1B, CCR1, STIL, BLM, C16ORF7, TPX2, PTI1, TCF3, CCNB1, DTX2, PIM1, ENSA, RCHY1, NFATC2IP, OBSL1 and MMP15.

9. The method according to claim 7 or 8, wherein the at least one gene probe is fixed on a microarray chip.

10. The method according to any of claims 7 to 9, wherein the measurement of gene expression level is performed by a microarray or quantitative reverse transcriptase polymerase chain reaction (quantitative RT-PCR).

11. The method according to claim 2 or 3, wherein the first pre-determined reference is a score of 31.

12. The method according to any of claims 4 to 6, wherein the second pre-determined reference is a score of 21.

13. The method according to claim 5 or 6, wherein the third pre-determined reference is a score of 44.

14. The method according to any of the preceding claims, wherein the subject has one of the following conditions: zero nodes, one to three positive nodes, and more than three positive nodes.

15. A method according to any of claims 1 to 14, wherein the predicted likelihood of LRR and/or distant metastasis is used to predict or determine the type of adjuvant treatment for the subject following mastectomy and/or breast conserving surgery.

16. A method for predicting the likelihood of locoregional recurrence (LRR) and/or distant metastasis in a subject with breast cancer following mastectomy and/or breast conserving surgery, comprising:

    i. measuring the expression level of a plurality of genes in a sample isolated from the subject; and

    ii. deriving a score based on the measured expression level of the plurality of genes using a predictive classification model, wherein the predictive classification model comprises at least one scoring algorithm;

wherein the plurality of genes is selected from a group consisting of: TRPV6, DDX39, BUB1B, CCR1, STIL, BLM, C16ORF7, TPX2, PTI1, TCF3, CCNB1, DTX2, PIM1, ENSA, RCHY1, NFATC2IP, OBSL1 and MMP15;

wherein the at least one scoring algorithm is selected from a group consisting of:

i. Score = TRPV6 + DDX39 + BUB1B + CCR1 + STIL + BLM + C16ORF7 + TPX2 + PTI1 + TCF3 + CCNB1 + DTX2 + PIM1 + ENSA + RCHY1 + NFATC2IP + OBSL1 + MMP15; wherein each gene counts as one point when the hazard ratio is <1, and the total score = 18;

ii. Score = TRPV6 + DDX39 + BUB1B + CCR1 + STIL + BLM + 2×C16ORF7 + TPX2 + PTI1 + TCF3 + CCNB1 + DTX2 + PIM1 + ENSA + RCHY1 + 2×NFATC2IP + OBSL1 + 2×MMP15; wherein when the hazard ratio is <1, each gene counts as one point if the genes

are examined by univariate analysis, and each gene counts as two points if the genes are examined by multivariate analysis, and so the total score = 21;

iii. Score = 3×TRPV6 + 5×DDX39 + 18×BUB1B + 4×CCR1 + 4×STIL + 7×BLM + 7×C16ORF7 + 4×PMI1 + 9×TPX2 + 8×PTI1 + 3×TCF3 + 7×CCNB1 + 2×DTX2 + 2×ENSA + 5×RCHY1 + 6×NFATC2IP + 2×OBSL1 + 6×MMP15; wherein the genes are examined by univariate analysis and the score of each gene is re-scaled according to its weighting, and so the total score = 102;

iv. Score = 2×TRPV6 + 2×DDX39 + 2×BUB1B + CCR1 + STIL + 2×BLM + 5×C16ORF7 + 3×PIM1 + 3×TPX2 + 5×PTI1 + TCF3 + CCNB1 + DTX2 + 2×ENSA + 3×RCHY1 + 4×NFATC2IP + OBSL1 + MMP15; wherein the genes are examined by multivariate analysis and the odds ratio of each gene is re-scaled to a score between 1 and 5, each gene counts as 1 point when the odds ratio is <1, and so the total score = 40; and

v. Score = 4×TRPV6 + 3×DDX39 + 8×BUB1B + CCR1 + STIL + 3×BLM + 11×C16ORF7 + 4×PIM1 + TPX2 + 2×PTI1 + 2×TCF3 + CCNB1 + DTX2 + 2×ENSA + 5×RCHY1 + 4×NFATC2IP + OBSL1 + 2×MMP15; wherein the genes are examined by multivariate analysis and the odds ratio of each gene is re-scaled to a score between 1 and 11, each gene counts as 1 point when the odds ratio is <1, and so the total score = 56; and

wherein the derived score provides an indication of the likelihood of LRR and/or the likelihood of distant metastasis in the subject.